

```

CC or send an email to license@sb-sib.ch).
CC EMBL: D63346; BAA09670.1; -
DR EMBL: X73325; CAA51751.1; -
DR HSP; P01730; 1WBR.
DR GO: GO:0042101; C/T-cell receptor complex; ISS.
DR GO: GO:0015026; Fc receptor activity; ISS.
DR GO: GO:0042289; Fc receptor class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045085; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO: GO:0030217; P-T-cell differentiation; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro: IPR000973; CD4 TCAG.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00447; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42
FT CARBOHYD 295 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 422
FT LIPID 422 422
FT CONFLICT 57 57
FT CONFLICT 91 91
FT CONFLICT 105 105
FT CONFLICT 113 113
FT CONFLICT 302 302
FT CONFLICT 349 349
SQ SEQUENCE 458 AA; 50905 MW; 751A9BA2CB83E16 CRC64;

Query Match 33.4%; Score 903; DB 1; Length 458;
Best Local Similarity 87.0%; Pred. No. 1.3e-50;
Matches 174; Conservative 14; Mismatches 12; Indels 0; Gaps 0;

```

```

DT 21-JUL-1986 (Rel. 01, Last sequence update)
DR 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig gamma chain C region.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6133520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RT F-I haplotype."
RL Immunogenetics 18:387-397(1983).
RN [2]
RP SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.B.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RT immunoglobulin G of different allotype.",
RL Biochem. J. 151:337-349(1975).
RN [3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steimetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RT heavy chain and identification of two genomic C gamma genes.",
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RN [4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Mole L.B., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RT immunoglobulin G."
RL Biochem. J. 116:249-259(1970).
RN [5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RT (In) Kiliander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
RL Stockholm (1967).
CC -1- MISCELLANEOUS; Ref.1 sequence has the D12 alleotypic marker,
CC 104-Thr, and the E14 marker, 185-Thr. Ref.3 has the D11 and E15
CC markers and Ref.5 the E15 marker.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL: M16426; AAA31289.1; -.
DR PIR: A91749; GHRR.
DR HSP; P01857; 1FC1.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_3.
DR SMART; SM00407; IGV; 2.
DR PROSITE; PS50835; IG LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON TER 1 1
FT DOMAIN 6 96
FT DOMAIN 114 213
FT DOMAIN 222 318
FT VARIANT 104 104
FT VARIANT 185 185
FT CONFLICT 48 48
FT CONFLICT 71 71

```

FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
 FT CONFLICT 173 173 N -> D (IN REF. 5).
 FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
 FT CONFLICT 201 201 N -> D (IN REF. 5).
 FT CONFLICT 218 218 Q -> E (IN REF. 5).
 FT CONFLICT 233 233 E -> Q (IN REF. 5).
 FT CONFLICT 246 246 N -> D (IN REF. 5).
 FT CONFLICT 256 256 E -> G (IN REF. 5).
 FT CONFLICT 260 260 N -> D (IN REF. 5).
 FT CONFLICT 266 266 N -> D (IN REF. 5).
 FT CONFLICT 280 280 Y -> W (IN REF. 5).
 FT CONFLICT 284 284 N -> S (IN REF. 5).
 SQ SEQUENCE 323 AA; 35404 MW; 6958AA118D579A8B CRC64;

Query Match 33.3%; Score 899; DB 1; Length 323;
 Best Local Similarity 61.0%; Pred. No. 1.5e-50;
 Matches 178; Conservative 34; Mismatches 64; Indels 16; Gaps 5;

QY 142 TLESPGSSPSVQCRPRGKNIQGGKTLVSQLELDQSGTWTCVYLQNKVFEKIDIVP 201
 DB 45 TLTVNGVTRFSPVQSS-----GLYSLSSVSVTSSSQPTVCNVA--HPATNTKVD--- 92
 QY 202 CPAPRPSKCDKHTHC--PELLGSPVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVK 259
 DB 93 -KTVAPSTCSKP-TCPPPELLGGPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVK 150
 QY 260 FNMVYGVGVHNAKTKPREEQVNSTYRVVSVLTVLHODMLNGKEKCKVSNKALPAPRIEK 319
 DB 151 FTWYINNQVTRARPLAEQPNSTIRVYSTLPITHQMLNGKEKCKVSNKALPAPRIEK 210
 QY 320 TISKAKGPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNQOPENNYKT 379
 DB 211 TISKARQPLEBPKVYTMGPPELSSRSVSLTCMNGFYPSPDISVEMEKNGKAEDNYKT 270
 QY 380 PVLVDSGSEFLXSKLTVDKSRMOQGNFSCSVHMEALHNHYTQKSLSLSG 431
 DB 271 PAVLDSDGSFLYLNKLVSPTSEWQGDVFTCSVMHEALHNHYTQKSLSRSG 322

RESULT 15

GC2_CAVPO STANDARD; PRT; 329 AA.
 AC P01862;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig gamma-2 chain C region.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognathii; Caviidae; Cavia.
 OC NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE OF 1-3.
 RA Trischmann T.M.;
 RL Submitted (Apr-1975) to the PIR data bank.
 RN [2]
 RP SEQUENCE OF 4-68.
 RX MEDLINE=71058471; PubMed=5538606;
 RA Birschtein B.K., Hussein O.Z., Cebra J.J.;
 RT "Structure of heavy chain from strain 13 guinea pig
 RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the
 RT half-cysteine joining heavy and light chains.";
 RL Biochemistry 10:18-25(1971).
 RN [3]
 RP SEQUENCE OF 69-133 AND 312-329.
 RX MEDLINE=71058486; PubMed=5538616;
 RA Turner K.J., Cebra J.J.;
 RT "Structure of heavy chain from strain 13 guinea pig
 RT immunoglobulin-G(2). 11. Amino acid sequence of the carboxyl-terminal
 RT and hinge region cyanoen bromide fragments.";
 RL Biochemistry 10:9-17(1971).
 RN [4]
 RP SEQUENCE OF 134-226.

RX MEDLINE=75036072; PubMed=4429665;
 RA Tracey D.E., Cebra J.J.;
 RT "Primary structure of the CH2 homology region from guinea pig IgG2
 RT antibodies.";
 RL Biochemistry 13:4796-4803(1974).
 RN [5]
 RP SEQUENCE OF 227-311.
 RX MEDLINE=75036073; PubMed=4609467;
 RA Trischmann T.M., Cebra J.J.;
 RT "Primary structure of the CH3 homology region from guinea pig IgG2
 RT antibodies.";
 RL Biochemistry 13:4804-4811(1974).
 RN [6]
 RP DISULFIDE BONDS.
 RX MEDLINE=71058474; PubMed=4922544;
 RA Oliveira B., Lamm M.E.;
 RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";
 RL Biochemistry 10:26-31(1971).
 CC -1- MISCELLANEOUS: This chain was isolated from pooled serum of strain
 CC 13 inbred guinea pigs.
 DR PIR: A94553; G2GP.
 DR HSSP; P01842; 7RAB.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig; 2.
 DR SMART; SM00407; IgC1; 2.
 DR PROSITE; PS50835; IG_LIKE; 3.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; glycoprotein.
 FT NOV TER 1
 FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 28 79
 FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 142 202
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
 FT DISULFID 248 308
 SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 33.1%; Score 894.5; DB 1; Length 329;
 Best Local Similarity 62.5%; Pred. No. 2.9e-50;
 Matches 172; Conservative 34; Mismatches 60; Indels 9; Gaps 4;

QY 163 IQGKTLVSQLELDQSGTWTCVYLQ--NQKXVEFKIDIVCPAPRPSKCDKHTHC--PE 218
 DB 58 LQSGLYSLTSMVTVVPSQKATCVNAPASSTKVDKTVETPRTPEBCTCPK--CPPE 114
 QY 219 LLAGPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFNMVYGVGVHNAKTKPRE 278
 DB 115 NLGSPSVFIRPPKPKDTLMISLTPRYTCVVDVSHEDPEVQFTWVNNKPVGNATETPRV 174
 QY 279 EQYNSTYRVVSVLTVLHODMLNGKEKCKVSNKALPAPRIEKTISKAKGPREPOVYTLPP 338
 DB 175 EGYVTFPRVSVLPICQHDMLRGKEFKCKVYNKALPAPRIEKTISKAKGPREPVPYTLPP 234
 QY 339 SRDELTKNOVSLTCLVKGFPSPDIAVWESNQOP--ENNYKTTPPVLDSDGSEFLYSKLT 396
 DB 235 SRDLSSKSVSVTCLINFPADIVHWMASNRVPSGKEKKNTPPIEDADGSFLYSLKLT 294
 QY 397 VDKSRMOQGNFSCSVHMEALHNHYTQKSLSLSG 431
 DB 295 VDKSAMQGTIVTCSVMHEALHNHYTQKSLSRSG 329

RESULT 16

CD4_CERAE STANDARD; PRT; 458 AA.
 AC O08138; O02805; O72593; Q28217;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)

T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).

Cercopithecus aethiops (Green monkey) (Grivet).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Cercopithecus.

NCBI_TaxID=9534;

SEQUENCE FROM N.A.

Hashimoto O., Tatemura M.;

"Molecular cloning and expression of african green monkey CD4.";

Submitted (Jul-1996) to the EMBL/GenBank/DBJ databases.

SEQUENCE OF 28-424 FROM N.A.

TISSUE=Blood;

MEDLINE=93049640; PubMed1425921;

Pomgaard A., Hirsch V.M., Johnson P.R.;

"Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus.";

Eur. J. Immunol. 22:2973-2981 (1992).

SEQUENCE OF 28-424 FROM N.A.

TISSUE=Peripheral blood;

MEDLINE=98017879; PubMed=9379478;

Pomgaard A., Mueller-Trutwin M.C., Diop O., Hansen J., Mathiot C., Corbet S., Barre-Sinoussi F., Allan J.S.;

"Relation between phylogeny of African green monkey CD4 genes and their respective simian immunodeficiency virus genes.";

J. Med. Primatol. 26:120-128 (1997).

SEQUENCE OF 107-192 FROM N.A.

MEDLINE=98320644; PubMed=9656488;

Harris E.E., Disocell T.R.;

"Nuclear gene trees and the phylogenetic relationships of the mangabeys (primates: Papionini).";

Mol. Biol. Evol. 15:892-900 (1998).

- FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.

- SUBUNIT: Associates with p56-lck (By similarity).

- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.

- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D86589; BAA13132.1; -

EMBL; X73322; CAA51748.1; -

EMBL; AF001226; AAB60873.1; -

EMBL; AF001228; AAB60873.1; -

EMBL; AF057380; AAC25124.1; -

HSSP; P01730; 1MIQ.

GO; GO:0042101; C-T-cell receptor complex; ISS.

GO; GO:0015026; F:coreceptor activity; ISS.

GO; GO:0042289; F:MHC class II protein binding; ISS.

GO; GO:0006955; P:immune response; ISS.

GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.

GO; GO:0030217; P:T-cell differentiation; ISS.

GO; GO:0045058; P:T-cell selection; ISS.

GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.

InterPro; IPR000973; CD4_TCRg.

InterPro; IPR007110; I9_Like.

InterPro; IPR003596; I9_V.

InterPro; IPR00047; I9_2.

InterPro; IPR00692; CD4TCANTIGEN.

PRINTS; SM00406; IGV; 1.

SMART; SM00406; IGV; 1.

PROSITE; PS50835; IG LIKE; 1.

Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell; Immune response; Repeat; Signal; Lipoprotein; Palmitate.

FT SIGNAL 1 25

FT CHAIN 26 458

FT DOMAIN 26 396

FT TRANSMEM 397 418

FT DOMAIN 419 458

FT DOMAIN 26 125

FT DOMAIN 126 203

FT DOMAIN 204 317

FT DOMAIN 318 374

FT CARBOHYD 42 42

FT CARBOHYD 281 281

FT CARBOHYD 296 296

FT CARBOHYD 325 325

FT DISULFID 41 109

FT DISULFID 155 184

FT DISULFID 328 370

FT LIPID 419 419

FT LIPID 422 422

FT CONFLICT 46 46

FT CONFLICT 59 59

FT CONFLICT 115 115

FT CONFLICT 165 165

FT CONFLICT 200 200

FT CONFLICT 227 227

FT CONFLICT 271 271

FT CONFLICT 281 281

SEQUENCE 458 AA; 51158 MW; FC523D2EDD1F72E7 CRC64;

Query Match 32.8%; Score 885; DB 1; Length 458;

Beet Local Similarity 85.5%; Pred. No. 1.8e-49;

Matches 171; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

1 MNRGVPPHLLLVLTQALLPAPATQGNKVLLKGGDTVELTCTASQKSIQPFMKNSNOIK 60

1 MNMGIPFPHLLLVLTQALLPAPVTOGKVVLLGKGGDTVELTCTASQKTTQPFMKNSNOIK 60

61 ILGNQSFVLTGPKSLNDRADSRSLMDQGNPPLIKLKIKEDSTTYICEVEDQKEVQL 120

61 ILGKQSFVLTGKSSSLDRIDSRKSLMDQGFSMIIKLIKIEDSEYTYICEVENKKEVEL 120

121 LVFGLTANSDPHLQGGSLTTLTLESPGSSPSVQCRSGRKNIOGKGLTSQLELDQSG 180

121 LVFGLTANSDPHLQGGSLTTLTLESPGSSPSVQCRSPRGKNIQGGRTLSVPQLERQDSG 180

181 TWTCTVLQNOKKVEPKIDIV 200

181 TWTCTVSQDQNTVERPKIDIM 200

RESULT 17

GCM MOUSE STANDARD; PRT; 405 AA.

AC P01867;

DT 21-VUL-1986 (Rel. 01, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE IG gamma-2b chain C region, membrane-bound form.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

SEQUENCE OF 335-405 FROM N.A.

MEDLINE=82222190; PubMed=6283537;

Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;

"Nucleotide sequences of gene segments encoding membrane domains of immunoglobulin gamma chains.";

Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627 (1982).

SEQUENCE OF 335-378 FROM N.A.

```

RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
RT immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
CC
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Membrane-bound;
CC IsoId=PI01867-1; Sequence=Displayed;
CC Name=Secreted;
CC IsoId=PI01866-1; Sequence=External;
CC Note=May be the major isoform;
CC -1- PTM: DISULFIDE BONDS BY HOMOLOG WITH OTHER IG GAMMA CHAINS.
CC MISCELLANEOUS: The sequence of residues 1-335 is assumed to be
CC identical with the corresponding region of the secreted form.
CC -1- MISCELLANEOUS: The a allele sequence is shown.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, J00462; AAB59659.1; ALT_INIT.
CC PIR, C02154; G2MSBM.
CC PDB, 1C1C; 1I-MAR-03.
CC MGD; MGI:96445; Igh-3.
CC InterPro: IPR007110; IG-like.
CC InterPro: IPR003597; IG_c1.
CC InterPro: IPR003006; IG_MHC.
CC Pfam: PF00047; Ig; 3.
CC SMART; SM00407; IgC1; 2.
CC PROSITE; PS50835; IG-LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Transmembrane;
CC Alternative splicing; 3d-structure; Repeat.
CC
CC FT NON_TER 1
CC FT DOMAIN 6
CC FT DOMAIN 127
CC FT DOMAIN 225
CC FT DISULFID 15
CC FT DISULFID 27
CC FT DISULFID 109
CC FT DISULFID 112
CC FT DISULFID 115
CC FT DISULFID 118
CC FT DISULFID 150
CC FT DISULFID 210
CC FT DISULFID 256
CC FT TRANSMEM 352
CC FT DOMAIN 370
CC SQ SEQUENCE 405 AA; 44330 MW; 89B3CF0A9B6D49FA CRC64;

Query Match 32.5%; Score 877.5; DB 1; Length 405;
Best Local Similarity 51.7%; Pred. No. 4.7e-49;
Matches 178; Conservative 52; Mismatches 75; Indels 39; Gaps 9;

OY 133 LIQG--OSLTLTLESPGSS-----PSVQCRSPRGNKIOGGKTLVSQLELDSDGT-- 182
DB LVKGYFPESVTVTNGSLSSSVHTFPAL-----LQSG-LVTMSSTVPSSTWPS 77
OY 183 ---TCTVQAO--NQKVEFKID-----IVPCPAEPKSCDKTHTC--PELLGSPVLF 228
DB QTVTSVAHPASSTVDDKLEPSGPISTINPCP-----PCKECHCPAPNLEGGSPVIF 132
OY 229 PPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAKTKRREQDYNSTYVV 288
DB 133 PPNIDVLMISTLPKVTLCVVVDVSDDDPQVLSMFWNVEVHTAQTQTHREDYNSITIKV 192
OY 289 SVLTIVLHODMNGKEYKCKVSNKALPAPIEKTISKAKGQPREPPQVYTLPPSRDELTKNQV 348

```

```

DB 193 STLFIQHODMNGSGKEFKCKVNNKOLPSPIBRTISKIGLVAPQVILPPAEOLSRKDV 252
OY 349 SLTCLVGFPPSDAVAVMENSGPENNYKTPPLTJDSGSEFLYSKTLVDSRMOQGVNF 408
DB 253 SLTCLVGFPPGSDISVETNSGHTENYKOTAPVLDGSGFYISKNNMKTISKWEKTDSE 312
OY 409 SCNVHBAALNHYTKSLSPGLQDPECAEADGSLDGLMTT 452
DB 313 SCNVHBAALNHYTKSLSPGLQDPECAEADGSLDGLMTT 356

RESULT 18
ID GC3 MOUSE STANDARD; PRT; 329 AA.
AC P22436.
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region, secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blatter F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBL J. 3:2041-2046(1984).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, J00451; -; NOT_ANNOTATED_CDS.
CC PIR, B02156; G3MSC.
CC HSSP; P01857; 1FC1.
CC InterPro: IPR007110; IG-like.
CC InterPro: IPR003597; IG_c1.
CC InterPro: IPR003006; IG_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00407; IgC1; 2.
CC PROSITE; PS50835; IG-LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
CC Transmembrane; Alternative splicing.
CC
CC FT NON_TER 1
CC FT DOMAIN 1
CC FT DOMAIN 98
CC FT DOMAIN 114
CC FT DOMAIN 224
CC SQ SEQUENCE 329 AA; 36628 MW; F45827174182BAD6 CRC64;

Query Match 31.7%; Score 857.5; DB 1; Length 329;
Best Local Similarity 59.6%; Pred. No. 6.8e-48;
Matches 162; Conservative 38; Mismatches 57; Indels 15; Gaps 4;

OY 170 SVSGLLELDSDGT-----TCTVLOKQKVEFKID---VPCPAEPKSCDKTHTC--EL 219
DB 62 SLSSLVTVPSSTWPSQTVICNVAPASKTELIRKIEPRIRKPSYPPS-----SCPGNI 116
OY 220 LGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAKTKRREQ 279
DB 117 LGSPSVIFPPPKKDTLMISTLPKVTLCVVVDVSDDDPQVLSMFWNVEVHTAQTQTHRED 176
OY 280 QNSTYRVSVLTVLHODMNGKEYKCKVSNKALPAPIEKTISKAKGQPREPPQVYTLPPS 339

```


Db 177 QVNSTFRVVSALPIQHDMWCKEFCVKVNNKALPAPIERTISKPKRAQTPQVYTIPTPP 236
 Qy 340 RDELTKNQVSLTCLVGVPSDIAVEMESNGOPENNYKTPPYLDSGSEFLYSKLTVDK 399
 Db 237 REOMSGKKKSLTCLVNFSEALSVEMERNGELQDYKTPPYLDSGTYFLYSKLTVD 296
 Qy 400 SRWQGNVSCSVMEALHNHYTQKSLSPG 431
 Db 297 DSWLGEIFTCVSVEALHNHYTQKSLSPG 328

RESULT 19

GCB_RAT ID GCB_RAT STANDARD; PRT; 333 AA.
 AC P20761;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE IG gamma-2B chain C region.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89232738; PubMed=3149946;
 RA Bruggemann M.;
 RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family."
 RL Gene 74:473-482(1988).
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
 DR HSSP; P01842; 7FAB.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003597; IG-cl.
 DR InterPro: IPR003006; IG_MHC.
 DR Pfam; PF00047; Ig1; 3.
 DR SMART; SM00407; IG1; 2.
 DR PROSITE; PSS0835; IG LIKE; 3.
 DR PROSITE; PSS0290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
 FT NON_TER 1 1
 FT DOMAIN 6 96 IG-LIKE 1.
 FT DOMAIN 124 223 IG-LIKE 2.
 FT DOMAIN 232 328 IG-LIKE 3.
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 80 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 147 207
 FT DISULFID 253 311
 SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 30.8%; Score 833; DB 1; Length 333;
 Best Local Similarity 58.6%; Pred. No. 2, 5e-46;
 Matches 157; Conservative 38; Mismatches 53; Indels 20; Gaps 5;

Qy 179 SGTW-----TCTVLO--NOKVEFKIDI-----VCPAPEPKSCDKTHTC--PELLGSP 223
 Db 70 SSTWPSQTYTCNVAVHAPASSTKVDKVERNRNGIGHKCP-----TCTCHKCPVELLGGP 124
 Qy 224 SYFLPFPKPKDTLMISTREPEVTCVVVDVSHEDPEVKFNMYVDQGVFNNAKTKPREEQYN 283
 Db 125 SYFLPFPKPKDTLLISQNAKVTCTCVVDVSEEDVQSFVFNWVEHTAQTQTPREEQYN 184
 Qy 284 TRVAVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDEL 343
 Db 185 TRVAVSALPIQHDMWCKEFCVKVNNKALPAPIEKTISKPKRAQTPQVYTIPTPP 244
 Qy 344 TKNQVSLTCLVGVPSDIAVEMESNGOPENNYKTPPYLDSGSEFLYSKLTVDKSRWQ 403
 Db 245 TEGTVSLTCLTSGFLPNDIGVETWSNGHTEKYNKTEPYWDSGSEFLYSKLTVDKSRW 304

Qy 404 QGNVSCSVMEALHNHYTQKSLSPG 431
 Db 305 SRAPVCSVMEALHNHYTQKSLSPG 332

RESULT 20

GCL_RAT ID GCL_RAT STANDARD; PRT; 326 AA.
 AC P20759;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG gamma-1 chain C region.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89232738; PubMed=3149946;
 RA Bruggemann M.;
 RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family."
 RL Gene 74:473-482(1988).
 CC PIR; PS0017; PS0017.
 DR HSSP; P01842; 7FAB.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003597; IG-cl.
 DR InterPro: IPR003006; IG_MHC.
 DR Pfam; PF00047; Ig1; 3.
 DR SMART; SM00407; IG1; 2.
 DR PROSITE; PSS0835; IG LIKE; 3.
 DR PROSITE; PSS0290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
 FT NON_TER 1 1
 FT DOMAIN 1 97 CH1.
 FT DOMAIN 98 112 HINGE.
 FT DOMAIN 113 219 CH2.
 FT DOMAIN 220 326 CH3.
 FT DISULFID 27 82
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 140 200
 FT DISULFID 246 304
 FT CARBOHYD 176 176
 SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;
 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 30.5%; Score 823.5; DB 1; Length 326;
 Best Local Similarity 50.0%; Pred. No. 1e-45;
 Matches 163; Conservative 53; Mismatches 55; Indels 55; Gaps 9;

Qy 133 ILQG---QSLTTLTSPGSS-----PSVQCRSPGKNIQGGKTLVSQLELDQSGTW-- 182
 Db 28 LVKGFPEPEVYTNMGSALSSGHTFPAY-----LQSLTYLTSSVTV--PSTWMS 77
 Qy 183 ---TCTVLO--NOKVEFKIDI--VCPAPEPKSCDKTHTCPELLG-----PSV 225
 Db 78 QTVTCNVAVHAPASSTKVDK-----VPRNC-----GDDCKPCICTGSEVSV 119
 Qy 226 FLFPKPKDTLMISTREPEVTCVVVDVSHEDPEVKFNMYVDQGVFNNAKTKPREEQYN 285
 Db 120 FLFPKPKDVLITLTPKTCVVVDISQDDPEVHFSWFDVDEVHTAQTQTPREEQYN 179
 Qy 286 RVAVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDEL 345
 Db 180 RVAVSALPIQHDMWCKEFCVKVNNKALPAPIEKTISKPKRAQTPQVYTIPTPP 244
 Qy 346 NOVSLTCLVGVPSDIAVEMESNGOPENNYKTPPYLDSGSEFLYSKLTVDKSRWQ 405
 Db 240 NEVSITCWKGFYPPDIYVEMQNGQPOBNYKNTPTMDTDSYFLYSKLTVDKSRWQ 299

QY 406 NVFSCVHMHANHYTOKSLSLSPG 431
 DB 300 NTFCTSVLHEGLHNHHTKSLSHSPG 325

RESULT 21

GCC_RAT STANDARD; PRT; 329 AA.
 ID_GCC_RAT
 AC P20762;
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig gamma-2C chain C region.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88166903; PubMed=3127222;
 RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
 RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
 RT region cDNA: extensive homology to mouse gamma 3.";
 RL Eur. J. Immunol. 18:317-319(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X07189; CAA30169.1; -
 DR PIR: S00847; S00847.
 DR HSR: F01842; 7PAB.
 DR InterPro: IPR007110; IG-1like.
 DR InterPro: IPR003597; IG_C1.
 DR InterPro: IPR003006; IG_MHC.
 DR Pfam: PF00047; Ig_2.
 DR SMART: SM00407; IGc1_2.
 DR PROSITE: PS50835; IG_LIKE; 3.
 DR PROSITE: PS00290; IG_MHC; 1.
 DR Immunoglobulin domain; Immunoglobulin C region.
 FT NON TER 1 1
 FT DOMAIN 1 97 CH1.
 FT 98 113 HINGE.
 FT DOMAIN 114 222 CH2.
 FT 223 329 CH3.
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 82
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 143 203
 FT DISULFID 249 307
 SQ SEQUENCE 329 AA; 36571 MW; 5PCD7B7933850773 CRC64;

Query Match 30.5%; Score 823; DB 1; Length 329;
 Best Local Similarity 58.5%; Pred. No. 1,1e-45;
 Matches 151; Conservative 45; Mismatches 56; Indels 6; Gaps 2;

QY 179 SGTW-----TCTVLNOKKVEFKIDIVCPAPEPKSCDKHTTCEPLLGSPVFLPFPKPK 233
 DB 72 SSTWSQVTCVAVHAPATKSNLIKRIEP-RRKRPAPPDIDCSGDNLRSPVETFPKPK 130
 QY 234 DTLMSIRPEVTCVAVDVSHDEPEVKFMVYDVGEVHNAKTKPREQVNSYRVVSVLTV 293
 DB 131 DILMITLTPKVTQVAVDVSEEDPDVQFSWFDNVAVFTAQIQPHBEQLNGFRVSTLHI 190
 QY 234 LHQDLNKEKYCKSKNSKALPAPIKTKSKAGQREQVYTLTPSPBELTKNQVSLTCL 353
 DB 191 QHQMWSGKEPKCKVNNKDLPSPIKTKISKPKGKARTQVYTIPTPPRQMSKNKSLTCLM 250

QY 354 VKGYPDIADVEMESNQPPNNYKTPPVLDSDGSPFLYSKLTVDKSRMOGNVSCSV 413
 DB 251 VTSFYPASISVEMERNELGDDYKNTLPVLDSDGSPFLYSKLTVDTSRMRKGDYTCV 310

QY 414 HEALHNHYTOKSLSLSPG 431
 DB 311 HEALHNHYTOKSLSLSPG 328

RESULT 22

GCL_MOUSE STANDARD; PRT; 324 AA.
 ID_GCL_MOUSE
 AC P01868;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ig gamma-2I chain C region secreted form.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80045036; PubMed=115593;
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 RT gamma 1 chain gene.";
 RL Cell 18:559-568(1979).
 RN [2]
 RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
 RX MEDLINE=8020559; PubMed=6769752;
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
 RT cloned in a bacterial plasmid.";
 RL Gene 9:87-97(1980).
 RN [3]
 RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=80012837; PubMed=113776;
 RA Rogers J., Clarke P., Salsler W.;
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
 RT heavy chain.";
 RL Nucleic Acids Res. 6:3305-3321(1979).
 RN [4]
 RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=78242288; PubMed=98524;
 RA Adegbo K.;
 RT "Evolution of immunoglobulin subclasses. Primary structure of a
 RT murine myeloma gamma1 chain.";
 RL J. Biol. Chem. 253:6068-6075(1978).
 RN [5]
 RP DISULFIDE BONDS (MOPC 21).
 RX MEDLINE=73008889; PubMed=5073237;
 RA Svasti J., Milstein C.;
 RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
 RL Biochem. J. 126:837-850(1972).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: Secreted.
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Secreted;
 CC Isoform=P01868-1; Sequence=Displayed;
 CC Note=May be the major isoform;
 CC Name=Membrane-bound;
 CC Isoform=P01869-1; Sequence=External;

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR EMBL; V00793; CAA24175.1; -
DR EMBL; V00795; CAA24176.1; -
DR PIR; A02159; GIMS.
DR GlycosylatedB; P01868; -
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR007110; Ig-1Ike.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS50835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
KM
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 HINGE.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 118 174
FT CARBOHYD 174 174
FT FTID=CAR_000055.
FT DISULFID 244 302
FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F3C93 CRC64;

Query Match 30.3%; Score 818.5; DB 1; Length 324;
Best Local Similarity 51.8%; Pred. No. 2.1e-45;
Matches 156; Conservative 50; Mismatches 60; Indels 35; Gaps 5;

QY 133 LLQGSGLTLTLSPSSPSVOCSPRGNIOGKTLVSQLELDPSTGTCVLAQ--NQ 190
DB 56 VLQSLYLTLSSSVTVPSSP-----RPSEVTCNVNHPASS 90
QY 191 KAVEEKIDVPCPAPEPKSCDKTHCPCLGSPVFLPPPKDTLMTSRPEVTCVVD 250
DB 91 TKVDKKIVPRDCGC-KPCIC-----TVPEV---SSVFIFPPKEDVLTITLTPKVCVVD 142
QY 251 VSHEDPEVKFNMYVGVGVNAKTKPREQVNSTYRVVSVLTVLHODMNGEKVKCYSN 310
DB 143 ISKDPDEVQFSMFVDVDEVHTAQOTQREQFNSTFVSSELTPIHODMNGEKPKCRVNS 202
QY 311 KALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVGFYPSTDAVWESNG 370
DB 203 AAFPAPIEKTISKTKGRPAKAPQVYITPPKKGMAKDQSLTCLMTIDFPEEDITVEMQNG 262
QY 371 QPENNYKTTTPVLDSDGSEFLYSKLTVDKSRNQGVNPSGVMEALNNHYTQKSLSLSP 430
DB 263 QPAENYKNTQPIPMNTNGSYFYVSKLVNQSNWEAGNTFTCSYLAHGLNHNHTKSLSSHP 322
QY 431 G 431
DB 323 G 323

```

```

DE Ig gamma-2a chain C region, A allele.
OC Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RL Balb/c gamma 2a heavy chain messenger RNA.";
RN Nucleic Acids Res. 8:3143-3155(1980).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=81189876; PubMed=6262729;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
and evolution of heavy chain genes: further evidence for intervening
RT sequence-mediated domain transfer.";
RN Nucleic Acids Res. 9:1365-1381(1981).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Olio R., Auffray C., Morchamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
RT suggests that exons can be exchanged between genes in a multigenic
family.";
RN Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RN (4)
RP MYELOMA PROTEIN MOPC 173.
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
RT immunoglobulin: amino-acid sequence of the FC fragment. Implications
RT for the evolution of immunoglobulin structure and function.";
RN Eur. J. Biochem. 43:423-435(1974).
RN (5)
RP DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
RT immunoglobulin. Identification of the disulfide bridges.";
RN Eur. J. Biochem. 30:452-462(1972).
CC -!- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; V00798; CAA24178.1; -
DR PIR; A02152; G2MSA.
DR PDB; 1E4W; 12-JUL-01.
DR PDB; 1E4X; 12-JUL-01.
DR PDB; 1MN0; 06-MAY-99.
DR InterPro; IPR007110; Ig-1Ike.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS50835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; 3D-structure; Repeat.
FT NON_TER 1
FT DOMAIN 1 98 IG-LIKE 1.
FT DOMAIN 121 220 IG-LIKE 2.
FT DOMAIN 229 325 IG-LIKE 3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).

```

FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 144 204
 FT DISULFID 250 308
 FT MOD RES 330 330
 SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 29.8%; Score 805.5; DB 1; Length 330;
 Best Local Similarity 58.5%; Pred. No. 1.4e-44;
 Matches 155; Conservative 33; Mismatches 58; Indels 19; Gaps 5;

QY 179 SGTW-----TCTVLO--NOKKVEKID-----IVCPAPBPKSCDKTTCBELLGSPVF 226
 DB 72 SSTWPSOSITCVVAPASSTKVKDKIEBGRPIKPCP---PCKC-----PAPNLLGSPSVF 124
 QY 227 LPPPKKPTLMSRPETLCVVVDVSHDEPKFVWYDVGVVNAHAKTKPREEQNSYR 286
 DB 125 IFFPKIKVLMISLPVITCVVDVSEDDPVQISFVNVNVEVHTAQOTHREDNSTLR 184
 QY 287 VSVVLTVLHODMLNGKEYCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKN 346
 DB 185 VVSALPIQHODMWSKEFKCKVNNKDLPAPIERTISKSGSVRAQVYVLEPPEEMTKK 244
 QY 347 QVSLTCLVKGFPYSDIAVWESNGQPENNYKTPPVLDSDGSFLYSLKLTVDKSRWQGN 406
 DB 245 QVTLTCMTDFWEPEDIVYEMTNGKTELNYKTEFLVLDSDGSYFWYSKLRVKKWVERN 304
 QY 407 VFSGSVMEALHNHYTOKSLSLSPG 431
 DB 305 SYSCSVHEGLHNHTTYSFKTPG 329

RESULT 24

GCAB_MOUSE STANDARD; PRT; 335 AA.

AC P01864;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ig gamma-2A chain C region secreted form (B allele).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=82037861; PubMed=6170065;
 RA Schreier P.H., Botwell A.L.M., Mueller-Hill B., Baltimore D.;
 RT "Multiple differences between the nucleic acid sequences of the
 RT IgG2a and IgG2b alleles of the mouse."
 RL Proc. Natl. Acad. Sci. U.S.A. 78:4495-4499(1981).
 RN [2]
 RP SEQUENCE.

RX MEDLINE=82037777; PubMed=6794027;
 RA Dognin M.J., Lauwereys M., Strosberg A.D.;
 RT "Multiple amino acid substitutions between murine gamma 2a heavy
 RT chain Cc regions of Ig1a and Ig1b allotypic forms."
 RL Proc. Natl. Acad. Sci. U.S.A. 78:4031-4035(1981).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Secreted;
 CC Isoform=Pol1864-1; Sequence=Displayed;
 CC Note=Probably the major isoform.

CC Name=Membrane-bound;
 CC Isoform=Pol1865-1; Sequence=External;
 CC -1- MISCELLANEOUS: The sequence differs from that of the a allele,
 CC from BALB/c mice, at 15% of the positions.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

CC -----
 CC EMBL; J00479; -; NOT_ANNOTATED_CDS.
 DR PIR; A02153; G2MSAB.
 DR PDB; 1BOG; 23-JAN-99.
 DR PDB; 1HH6; 26-JAN-01.
 DR PDB; 1HH9; 24-JUL-03.
 DR PDB; 1HT6; 08-FEB-01.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003597; Ig c1.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig; 2.
 DR SMART; SM00407; IgC1; 2.
 DR PROSITE; PS50835; Ig_LIKE; 3.
 DR PROSITE; PS00290; Ig_MHC; 1.
 DR Immunoglobulin domain; Immunoglobulin C region; Alternative splicing;
 DR 3D-structure; Repeat.
 KW NON TER 1
 FT DOMAIN 6 98 IG-LIKE 1.
 FT DOMAIN 126 225 IG-LIKE 2.
 FT DOMAIN 234 330 IG-LIKE 3.
 SQ SEQUENCE 335 AA; 36596 MW; FA3382792CB813C6 CRC64;

Query Match 29.6%; Score 801; DB 1; Length 335;
 Best Local Similarity 50.2%; Pred. No. 2.8e-44;
 Matches 166; Conservative 52; Mismatches 77; Indels 36; Gaps 9;

QY 124 GLTANSDT--HLQG---QSLTTLLESPGSS-----PSVQCRSPRKNIGSGKTLVSQ 173
 DB 17 GTTGSVTLGCLVGYPEPVTILWNKSLSGVHTPAL-----LQSG-LYTLSS 66
 QY 174 LELDQSGTW-----TCTV-----LQNKVVEKIDIV--PCPAPBPKSCDKTTCBELL 220
 DB 67 SVTYSNTWPSQITTCVVAHPASSTKVKDKIEBGRPIYTONPCP---PQGRVPCAADLL 123
 QY 221 GGPSEVFPFPKPKDTLMSRPETLCVVVDVSHDEPKFVWYDVGVVNAHAKTKPREEQ 280
 DB 124 GGPSEVFPFPKPKDTLMSRPETLCVVVDVSHDEPKFVWYDVGVVNAHAKTKPREEQ 183
 QY 281 YNSTYRVSVYLVTHQDMLNGKEYCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSR 340
 DB 184 YNSTLRVVSALPIQHODMWSKEFKCKVNNKDLPAPIERTISKSGSVRAQVYVLEPPE 243
 QY 341 DELTKNOVSLTCLVKGFPYSDIAVWESNGQPENNYKTPPVLDSDGSFLYSLKLTVDKS 400
 DB 244 EEMTKKEFSLTCTMITGFLPAEIAVDWTSNGRTSQNYKNTATVLDSDGSYFWYSKLRVQKS 303
 QY 401 RMQGNVFSGSVMEALHNHYTOKSLSLSPG 431
 DB 304 TWERKSLFACSVVHEVHLNHLTKTISRSLG 334

RESULT 25

GCA_RAT STANDARD; PRT; 322 AA.

AC P20760;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ig gamma-2A chain C region.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 NC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89232738; PubMed=3149946;
 RA Brueggemann M.;

RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";

```

RL Gene 74:473-482(1988).
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M13804; AAA41376.1; ALU_INIT.
CC PIR: PS0019; PS0019.
CC HSSP: P01842; 7EAB.
CC InterPro: IPR007110; Ig-like.
CC InterPro: IPR003557; Ig_C1.
CC InterPro: IPR003006; Ig_MHC.
CC Pfam: PF00047; Ig_2.
CC SMART: SM00407; IgC1; 2.
CC PROSITE: PS50835; IG_LIKE; 3.
CC PROSITE: PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Repeat.
KW NON_TER
FT DOMAIN 1 98 IG-LIKE 1.
FT DOMAIN 115 212 IG-LIKE 2.
FT DOMAIN 221 317 IG-LIKE 3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 136 196
FT DISULFID 242 300
FT CARBOHYD 172 172
SQ SEQUENCE 322 AA; 35186 MW; E8EA136A9DE01EDB CRC64;
Query Match 29.3%; Score 792.5; DB 1; Length 322;
Best Local Similarity 51.0%; Pred. No. 9,4e-44;
Matches 160; Conservative 47; Mismatches 72; Indels 35; Gaps 8;
QY 133 LLOG--QSLTLTLESPSS-----PSVQCSPPKNGKIQGKTLTSLVQLQDSGTW-- 182
DB 28 LVKGYPEPEVTVYVNGALSSGVHTPPAV-----LQSGLYTLTSSVTV-PSSTWSS 77
QY 183 ----TCTVLO--NOKVEFIDIVPCPAPRPKSCDKHTHCEBLLGGSVSLFPKPKDTLM 237
DB 78 QAVTCAVAHPASTKVDKXIVPREC--NPCGTGSEV-----SSVPIFPKPTDVLTL 127
QY 238 ISRTPEVTGVVDSHEDEPEVKENMYVDGVEVNAATKPREQYNSTYRVSVLTVLHOD 297
DB 128 ITLTPKVTGVVVDISQNDPEVRSMVIDVEVHTAQTTHAPRKQSNSTLRSVSELPVHRD 187
QY 298 WLNKGEYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSHDELTKNQVSLTCLVKGF 357
DB 188 WLNKGTFFKCKVNSGAPAPIEKISKPEGTPRGPOVYTAAPREENTQGSVITCWVKG 247
QY 358 YPSDIAVESNGQPENNYKTPPVLDSDGSFPLYSKLTVDSKRWQGVFSGSVAREHL 417
DB 248 YPDIDYTEKRWKNGQPENNYKTPPTDSDGSYFLYSKLVNKKETWQGNFTTCSVLAHEL 307
QY 418 HNHYYTOKSLSPG 431
DB 308 HNHHTKSLSHSPG 321

```

```

GN CD4.
OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=95311;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA MEDLINE=93049640; PubMed=1425921;
RX Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X73328; CAAS1754.1; -.
CC EMBL: X73327; CAAS1753.1; -.
CC HSSP: P01730; 1WTO.
CC DR GO: GO:0042101; C: T-cell receptor complex; ISS.
CC DR GO: GO:0015026; F: coreceptor activity; ISS.
CC DR GO: GO:0042289; F: MHC class II protein binding; ISS.
CC DR GO: GO:0006955; P: immune response; ISS.
CC DR GO: GO:0045086; P: positive regulation of interleukin-2 biosyn. .; ISS.
CC DR GO: GO:0030217; P: cell differentiation; ISS.
CC DR GO: GO:0045058; P: T-cell selection; ISS.
CC DR GO: GO:0007169; P: transmembrane receptor protein tyrosine kin. .; ISS.
CC DR InterPro: IPR000973; CD4_TCRG.
CC DR InterPro: IPR007110; Ig-Like.
CC DR InterPro: IPR003596; Ig_V.
CC DR Pfam: PF00047; Ig_2.
CC DR PRINTS: PR00692; CD4TCANTIGEN.
CC DR SMART: SM00406; IG_V; 1.
CC DR PROSITE: PS50835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Lipoprotein; Palmitate.
FT DOMAIN 1 369
FT DOMAIN 370 391 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 392 >397 CYTOPLASMIC (POTENTIAL).
FT DOMAIN <1 98 IG-LIKE V-TYPE.
FT DOMAIN 99 176 IG-LIKE C2-TYPE 1.
FT DOMAIN 177 290 IG-LIKE C2-TYPE 2.
FT DOMAIN 291 347 IG-LIKE C2-TYPE 3.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 14 82 BY SIMILARITY.
FT DISULFID 128 157 BY SIMILARITY.
FT DISULFID 301 343 BY SIMILARITY.
FT LIPID 392 392 S-palmitoyl cysteine (By similarity).
FT LIPID 395 395 S-palmitoyl cysteine (By similarity).
FT VARIANT 20 20 MISSING.
FT VARIANT 43 43 T -> I.
FT VARIANT 86 86 N -> D.
FT VARIANT 96 96 F -> L.
FT VARIANT 173 173 V -> M.
FT VARIANT 316 316 R -> K.
FT NON_TER 397

```

```

SQ SEQUENCE 397 AA; 43926 MW; 8660B636D2DB38A7 CRC64;
Query Match 29.0%; Score 784; DB 1; Length 397;
Best Local Similarity 87.3%; Pred. No. 4.3e-43;
Matches 151; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 28 VLGGKGDVTELCTASQKKSIOFHWKNSNOIKILGNQGSFLTGPSKLNDRADSRSLW 87
DB 1 VLGGKGDVTELACNASQKSTQFHWKNSKQIKILGNQGSFLTQSSKLSRADSRSLW 60
QY 88 DQGNFPLIINKLIKEDSDTYICEVEDQKEEVOVLVFGITANSPTHLLOGOSLTLTLESPP 147
DB 61 DQGCFSMIINKLIKEDSEITYICEVENKKEVEVLVFGITANSPTHLLOGOSLTLTLESPP 120
QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTCTVLONOKKVEFKIDIV 200
DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTCTVLONOKKVEFKIDIV 173

RESULT 27
CD4_ERYPA STANDARD; PRT; 397 AA.
AC 008339; Rel. 31, Created)
DT 01-EB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3)
DE (Fragment).
OS Erythrocytes pates (Red guenon) (Husar).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Erythrocytes.
OX NCBI_TaxId=9538;
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus."
RL Eur. J. Immunol. 22:2973-2981(1992).
CC - FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC - SUBUNIT: Associates with p56-lck (By similarity).
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X73324; CA51750.1; -.
DR HSBP; P01730; I1T0.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR000973; CD4 TCR.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.

```

```

DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Lipoprotein; Palmitate.
FT NON_TER 1
FT DOMAIN 1 369
FT TIT 370
FT TIT 391
FT TIT 392
FT TIT 393
FT TIT 394
FT TIT 395
FT TIT 396
FT TIT 397
FT TIT 398
FT TIT 399
FT TIT 400
FT TIT 401
FT TIT 402
FT TIT 403
FT TIT 404
FT TIT 405
FT TIT 406
FT TIT 407
FT TIT 408
FT TIT 409
FT TIT 410
FT TIT 411
FT TIT 412
FT TIT 413
FT TIT 414
FT TIT 415
FT TIT 416
FT TIT 417
FT TIT 418
FT TIT 419
FT TIT 420
FT TIT 421
FT TIT 422
FT TIT 423
FT TIT 424
FT TIT 425
FT TIT 426
FT TIT 427
FT TIT 428
FT TIT 429
FT TIT 430
FT TIT 431
FT TIT 432
FT TIT 433
FT TIT 434
FT TIT 435
FT TIT 436
FT TIT 437
FT TIT 438
FT TIT 439
FT TIT 440
FT TIT 441
FT TIT 442
FT TIT 443
FT TIT 444
FT TIT 445
FT TIT 446
FT TIT 447
FT TIT 448
FT TIT 449
FT TIT 450
FT TIT 451
FT TIT 452
FT TIT 453
FT TIT 454
FT TIT 455
FT TIT 456
FT TIT 457
FT TIT 458
FT TIT 459
FT TIT 460
FT TIT 461
FT TIT 462
FT TIT 463
FT TIT 464
FT TIT 465
FT TIT 466
FT TIT 467
FT TIT 468
FT TIT 469
FT TIT 470
FT TIT 471
FT TIT 472
FT TIT 473
FT TIT 474
FT TIT 475
FT TIT 476
FT TIT 477
FT TIT 478
FT TIT 479
FT TIT 480
FT TIT 481
FT TIT 482
FT TIT 483
FT TIT 484
FT TIT 485
FT TIT 486
FT TIT 487
FT TIT 488
FT TIT 489
FT TIT 490
FT TIT 491
FT TIT 492
FT TIT 493
FT TIT 494
FT TIT 495
FT TIT 496
FT TIT 497
FT TIT 498
FT TIT 499
FT TIT 500
FT TIT 501
FT TIT 502
FT TIT 503
FT TIT 504
FT TIT 505
FT TIT 506
FT TIT 507
FT TIT 508
FT TIT 509
FT TIT 510
FT TIT 511
FT TIT 512
FT TIT 513
FT TIT 514
FT TIT 515
FT TIT 516
FT TIT 517
FT TIT 518
FT TIT 519
FT TIT 520
FT TIT 521
FT TIT 522
FT TIT 523
FT TIT 524
FT TIT 525
FT TIT 526
FT TIT 527
FT TIT 528
FT TIT 529
FT TIT 530
FT TIT 531
FT TIT 532
FT TIT 533
FT TIT 534
FT TIT 535
FT TIT 536
FT TIT 537
FT TIT 538
FT TIT 539
FT TIT 540
FT TIT 541
FT TIT 542
FT TIT 543
FT TIT 544
FT TIT 545
FT TIT 546
FT TIT 547
FT TIT 548
FT TIT 549
FT TIT 550
FT TIT 551
FT TIT 552
FT TIT 553
FT TIT 554
FT TIT 555
FT TIT 556
FT TIT 557
FT TIT 558
FT TIT 559
FT TIT 560
FT TIT 561
FT TIT 562
FT TIT 563
FT TIT 564
FT TIT 565
FT TIT 566
FT TIT 567
FT TIT 568
FT TIT 569
FT TIT 570
FT TIT 571
FT TIT 572
FT TIT 573
FT TIT 574
FT TIT 575
FT TIT 576
FT TIT 577
FT TIT 578
FT TIT 579
FT TIT 580
FT TIT 581
FT TIT 582
FT TIT 583
FT TIT 584
FT TIT 585
FT TIT 586
FT TIT 587
FT TIT 588
FT TIT 589
FT TIT 590
FT TIT 591
FT TIT 592
FT TIT 593
FT TIT 594
FT TIT 595
FT TIT 596
FT TIT 597
FT TIT 598
FT TIT 599
FT TIT 600
FT TIT 601
FT TIT 602
FT TIT 603
FT TIT 604
FT TIT 605
FT TIT 606
FT TIT 607
FT TIT 608
FT TIT 609
FT TIT 610
FT TIT 611
FT TIT 612
FT TIT 613
FT TIT 614
FT TIT 615
FT TIT 616
FT TIT 617
FT TIT 618
FT TIT 619
FT TIT 620
FT TIT 621
FT TIT 622
FT TIT 623
FT TIT 624
FT TIT 625
FT TIT 626
FT TIT 627
FT TIT 628
FT TIT 629
FT TIT 630
FT TIT 631
FT TIT 632
FT TIT 633
FT TIT 634
FT TIT 635
FT TIT 636
FT TIT 637
FT TIT 638
FT TIT 639
FT TIT 640
FT TIT 641
FT TIT 642
FT TIT 643
FT TIT 644
FT TIT 645
FT TIT 646
FT TIT 647
FT TIT 648
FT TIT 649
FT TIT 650
FT TIT 651
FT TIT 652
FT TIT 653
FT TIT 654
FT TIT 655
FT TIT 656
FT TIT 657
FT TIT 658
FT TIT 659
FT TIT 660
FT TIT 661
FT TIT 662
FT TIT 663
FT TIT 664
FT TIT 665
FT TIT 666
FT TIT 667
FT TIT 668
FT TIT 669
FT TIT 670
FT TIT 671
FT TIT 672
FT TIT 673
FT TIT 674
FT TIT 675
FT TIT 676
FT TIT 677
FT TIT 678
FT TIT 679
FT TIT 680
FT TIT 681
FT TIT 682
FT TIT 683
FT TIT 684
FT TIT 685
FT TIT 686
FT TIT 687
FT TIT 688
FT TIT 689
FT TIT 690
FT TIT 691
FT TIT 692
FT TIT 693
FT TIT 694
FT TIT 695
FT TIT 696
FT TIT 697
FT TIT 698
FT TIT 699
FT TIT 700
FT TIT 701
FT TIT 702
FT TIT 703
FT TIT 704
FT TIT 705
FT TIT 706
FT TIT 707
FT TIT 708
FT TIT 709
FT TIT 710
FT TIT 711
FT TIT 712
FT TIT 713
FT TIT 714
FT TIT 715
FT TIT 716
FT TIT 717
FT TIT 718
FT TIT 719
FT TIT 720
FT TIT 721
FT TIT 722
FT TIT 723
FT TIT 724
FT TIT 725
FT TIT 726
FT TIT 727
FT TIT 728
FT TIT 729
FT TIT 730
FT TIT 731
FT TIT 732
FT TIT 733
FT TIT 734
FT TIT 735
FT TIT 736
FT TIT 737
FT TIT 738
FT TIT 739
FT TIT 740
FT TIT 741
FT TIT 742
FT TIT 743
FT TIT 744
FT TIT 745
FT TIT 746
FT TIT 747
FT TIT 748
FT TIT 749
FT TIT 750
FT TIT 751
FT TIT 752
FT TIT 753
FT TIT 754
FT TIT 755
FT TIT 756
FT TIT 757
FT TIT 758
FT TIT 759
FT TIT 760
FT TIT 761
FT TIT 762
FT TIT 763
FT TIT 764
FT TIT 765
FT TIT 766
FT TIT 767
FT TIT 768
FT TIT 769
FT TIT 770
FT TIT 771
FT TIT 772
FT TIT 773
FT TIT 774
FT TIT 775
FT TIT 776
FT TIT 777
FT TIT 778
FT TIT 779
FT TIT 780
FT TIT 781
FT TIT 782
FT TIT 783
FT TIT 784
FT TIT 785
FT TIT 786
FT TIT 787
FT TIT 788
FT TIT 789
FT TIT 790
FT TIT 791
FT TIT 792
FT TIT 793
FT TIT 794
FT TIT 795
FT TIT 796
FT TIT 797
FT TIT 798
FT TIT 799
FT TIT 800
FT TIT 801
FT TIT 802
FT TIT 803
FT TIT 804
FT TIT 805
FT TIT 806
FT TIT 807
FT TIT 808
FT TIT 809
FT TIT 810
FT TIT 811
FT TIT 812
FT TIT 813
FT TIT 814
FT TIT 815
FT TIT 816
FT TIT 817
FT TIT 818
FT TIT 819
FT TIT 820
FT TIT 821
FT TIT 822
FT TIT 823
FT TIT 824
FT TIT 825
FT TIT 826
FT TIT 827
FT TIT 828
FT TIT 829
FT TIT 830
FT TIT 831
FT TIT 832
FT TIT 833
FT TIT 834
FT TIT 835
FT TIT 836
FT TIT 837
FT TIT 838
FT TIT 839
FT TIT 840
FT TIT 841
FT TIT 842
FT TIT 843
FT TIT 844
FT TIT 845
FT TIT 846
FT TIT 847
FT TIT 848
FT TIT 849
FT TIT 850
FT TIT 851
FT TIT 852
FT TIT 853
FT TIT 854
FT TIT 855
FT TIT 856
FT TIT 857
FT TIT 858
FT TIT 859
FT TIT 860
FT TIT 861
FT TIT 862
FT TIT 863
FT TIT 864
FT TIT 865
FT TIT 866
FT TIT 867
FT TIT 868
FT TIT 869
FT TIT 870
FT TIT 871
FT TIT 872
FT TIT 873
FT TIT 874
FT TIT 875
FT TIT 876
FT TIT 877
FT TIT 878
FT TIT 879
FT TIT 880
FT TIT 881
FT TIT 882
FT TIT 883
FT TIT 884
FT TIT 885
FT TIT 886
FT TIT 887
FT TIT 888
FT TIT 889
FT TIT 890
FT TIT 891
FT TIT 892
FT TIT 893
FT TIT 894
FT TIT 895
FT TIT 896
FT TIT 897
FT TIT 898
FT TIT 899
FT TIT 900
FT TIT 901
FT TIT 902
FT TIT 903
FT TIT 904
FT TIT 905
FT TIT 906
FT TIT 907
FT TIT 908
FT TIT 909
FT TIT 910
FT TIT 911
FT TIT 912
FT TIT 913
FT TIT 914
FT TIT 915
FT TIT 916
FT TIT 917
FT TIT 918
FT TIT 919
FT TIT 920
FT TIT 921
FT TIT 922
FT TIT 923
FT TIT 924
FT TIT 925
FT TIT 926
FT TIT 927
FT TIT 928
FT TIT 929
FT TIT 930
FT TIT 931
FT TIT 932
FT TIT 933
FT TIT 934
FT TIT 935
FT TIT 936
FT TIT 937
FT TIT 938
FT TIT 939
FT TIT 940
FT TIT 941
FT TIT 942
FT TIT 943
FT TIT 944
FT TIT 945
FT TIT 946
FT TIT 947
FT TIT 948
FT TIT 949
FT TIT 950
FT TIT 951
FT TIT 952
FT TIT 953
FT TIT 954
FT TIT 955
FT TIT 956
FT TIT 957
FT TIT 958
FT TIT 959
FT TIT 960
FT TIT 961
FT TIT 962
FT TIT 963
FT TIT 964
FT TIT 965
FT TIT 966
FT TIT 967
FT TIT 968
FT TIT 969
FT TIT 970
FT TIT 971
FT TIT 972
FT TIT 973
FT TIT 974
FT TIT 975
FT TIT 976
FT TIT 977
FT TIT 978
FT TIT 979
FT TIT 980
FT TIT 981
FT TIT 982
FT TIT 983
FT TIT 984
FT TIT 985
FT TIT 986
FT TIT 987
FT TIT 988
FT TIT 989
FT TIT 990
FT TIT 991
FT TIT 992
FT TIT 993
FT TIT 994
FT TIT 995
FT TIT 996
FT TIT 997
FT TIT 998
FT TIT 999
FT TIT 1000
FT TIT 1001
FT TIT 1002
FT TIT 1003
FT TIT 1004
FT TIT 1005
FT TIT 1006
FT TIT 1007
FT TIT 1008
FT TIT 1009
FT TIT 1010
FT TIT 1011
FT TIT 1012
FT TIT 1013
FT TIT 1014
FT TIT 1015
FT TIT 1016
FT TIT 1017
FT TIT 1018
FT TIT 1019
FT TIT 1020
FT TIT 1021
FT TIT 1022
FT TIT 1023
FT TIT 1024
FT TIT 1025
FT TIT 1026
FT TIT 1027
FT TIT 1028
FT TIT 1029
FT TIT 1030
FT TIT 1031
FT TIT 1032
FT TIT 1033
FT TIT 1034
FT TIT 1035
FT TIT 1036
FT TIT 1037
FT TIT 1038
FT TIT 1039
FT TIT 1040
FT TIT 1041
FT TIT 1042
FT TIT 1043
FT TIT 1044
FT TIT 1045
FT TIT 1046
FT TIT 1047
FT TIT 1048
FT TIT 1049
FT TIT 1050
FT TIT 1051
FT TIT 1052
FT TIT 1053
FT TIT 1054
FT TIT 1055
FT TIT 1056
FT TIT 1057
FT TIT 1058
FT TIT 1059
FT TIT 1060
FT TIT 1061
FT TIT 1062
FT TIT 1063
FT TIT 1064
FT TIT 1065
FT TIT 1066
FT TIT 1067
FT TIT 1068
FT TIT 1069
FT TIT 1070
FT TIT 1071
FT TIT 1072
FT TIT 1073
FT TIT 1074
FT TIT 1075
FT TIT 1076
FT TIT 1077
FT TIT 1078
FT TIT 1079
FT TIT 1080
FT TIT 1081
FT TIT 1082
FT TIT 1083
FT TIT 1084
FT TIT 1085
FT TIT 1086
FT TIT 1087
FT TIT 1088
FT TIT 1089
FT TIT 1090
FT TIT 1091
FT TIT 1092
FT TIT 1093
FT TIT 1094
FT TIT 1095
FT TIT 1096
FT TIT 1097
FT TIT 1098
FT TIT 1099
FT TIT 1100
FT TIT 1101
FT TIT 1102
FT TIT 1103
FT TIT 1104
FT TIT 1105
FT TIT 1106
FT TIT 1107
FT TIT 1108
FT TIT 1109
FT TIT 1110
FT TIT 1111
FT TIT 1112
FT TIT 1113
FT TIT 1114
FT TIT 1115
FT TIT 1116
FT TIT 1117
FT TIT 1118
FT TIT 1119
FT TIT 1120
FT TIT 1121
FT TIT 1122
FT TIT 1123
FT TIT 1124
FT TIT 1125
FT TIT 1126
FT TIT 1127
FT TIT 1128
FT TIT 1129
FT TIT 1130
FT TIT 1131
FT TIT 1132
FT TIT 1133
FT TIT 1134
FT TIT 1135
FT TIT 1136
FT TIT 1137
FT TIT 1138
FT TIT 1139
FT TIT 1140
FT TIT 1141
FT TIT 1142
FT TIT 1143
FT TIT 1144
FT TIT 1145
FT TIT 1146
FT TIT 1147
FT TIT 1148
FT TIT 1149
FT TIT 1150
FT TIT 1151
FT TIT 1152
FT TIT 1153
FT TIT 1154
FT TIT 1155
FT TIT 1156
FT TIT 1157
FT TIT 1158
FT TIT 1159
FT TIT 1160
FT TIT 1161
FT TIT 1162
FT TIT 1163
FT TIT 1164
FT TIT 1165
FT TIT 1166
FT TIT 1167
FT TIT 1168
FT TIT 1169
FT TIT 1170
FT TIT 1171
FT TIT 1172
FT TIT 1173
FT TIT 1174
FT TIT 1175
FT TIT 1176
FT TIT 1177
FT TIT 1178
FT TIT 1179
FT TIT 1180
FT TIT 1181
FT TIT 1182
FT TIT 1183
FT TIT 1184
FT TIT 1185
FT TIT 1186
FT TIT 1187
FT TIT 1188
FT TIT 1189
FT TIT 1190
FT TIT 1191
FT TIT 1192
FT TIT 1193
FT TIT 1194
FT TIT 1195
FT TIT 1196
FT TIT 1197
FT TIT 1198
FT TIT 1199
FT TIT 1200
FT TIT 1201
FT TIT 1202
FT TIT 1203
FT TIT 1204
FT TIT 1205
FT TIT 1206
FT TIT 1207
FT TIT 1208
FT TIT 1209
FT TIT 1210
FT TIT 1211
FT TIT 1212
FT TIT 1213
FT TIT 1214
FT TIT 1215
FT TIT 1216
FT TIT 1217
FT TIT 1218
FT TIT 1219
FT TIT 1220
FT TIT 1221
FT TIT 1222
FT TIT 1223
FT TIT 1224
FT TIT 1225
FT TIT 1226
FT TIT 1227
FT TIT 1228
FT TIT 1229
FT TIT 1230
FT TIT 1231
FT TIT 1232
FT TIT 1233
FT TIT 1234
FT TIT 1235
FT TIT 1236
FT TIT 1237
FT TIT 1238
FT TIT 1239
FT TIT 1240
FT TIT 1241
FT TIT 1242
FT TIT 1243
FT TIT 1244
FT TIT 1245
FT TIT 1246
FT TIT 1247
FT TIT 1248
FT TIT 1249
FT TIT 1250
FT TIT 1251
FT TIT 1252
FT TIT 1253
FT TIT 1254
FT TIT 1255
FT TIT 1256
FT TIT 1257
FT TIT 1258
FT TIT 1259
FT TIT 1260
FT TIT 1261
FT TIT 1262
FT TIT 1263
FT TIT 1264
FT TIT 1265
FT TIT 1266
FT TIT 1267
FT TIT 1268
FT TIT 1269
FT TIT 1270
FT TIT 1271
FT TIT 1272
FT TIT 1273
FT TIT 1274
FT TIT 1275
FT TIT 1276
FT TIT 1277
FT TIT 1278
FT TIT 1279
FT TIT 1280
FT TIT 1281
FT TIT 1282
FT TIT 1283
FT TIT 1284
FT TIT 1285
FT TIT 1286
FT TIT 1287
FT TIT 1288
FT TIT 1289
FT TIT 1290
FT TIT 1291
FT TIT 1292
FT TIT 1293
FT TIT 1294
FT TIT 1295
FT TIT 1296
FT TIT 1297
FT TIT 1298
FT TIT 1299
FT TIT 1300
FT TIT 1301
FT TIT 1302
FT TIT 1303
FT TIT 1304
FT TIT 1305
FT TIT 1306
FT TIT 1307
FT TIT 1308
FT TIT 1309
FT TIT 1310
FT TIT 1311
FT TIT 1312
FT TIT 1313
FT TIT 1314
FT TIT 1315
FT TIT 1316
FT TIT 1317
FT TIT 1318
FT TIT 1319
FT TIT 1320
FT TIT 1321
FT TIT 1322
FT TIT 1323
FT TIT 1324
FT TIT 1325
FT TIT 1326
FT TIT 1327
FT TIT 1328
FT TIT 1329
FT TIT 1330
FT TIT 1331
FT TIT 1332
FT TIT 1333
FT TIT 1334
FT TIT 1335
FT TIT 1336
FT TIT 1337
FT TIT 1338
FT TIT 1339
FT TIT 1340
FT TIT 1341
FT TIT 1342
FT TIT 1343
FT TIT 1344
FT TIT 1345
FT TIT 1346
FT TIT 1347
FT TIT 1348
FT TIT 1349
FT TIT 1350
FT TIT 1351
FT TIT 1352
FT TIT 1353
FT TIT 1354
FT TIT 1355
FT TIT 1356
FT TIT 1357
FT TIT 1358
FT TIT 1359
FT TIT 1360
FT TIT 1361
FT TIT 1362
FT TIT 1363
FT TIT 1364
FT TIT 1365
FT TIT 1366
FT TIT 1367
FT TIT 1368
FT TIT 1369
FT TIT 1370
FT TIT 1371
FT TIT 1372
FT TIT 1373
FT TIT 1374
FT TIT 1375
FT TIT 1376
FT TIT 1377
FT TIT 1378
FT TIT 1379
FT TIT 1380
FT TIT 1381
FT TIT 1382
FT TIT 1383
FT TIT 1384
FT TIT 1385
FT TIT 1386
FT TIT 1387
FT TIT 1388
FT TIT 1389
FT TIT 1390
FT TIT 1391
FT TIT 1392
FT TIT 1393
FT TIT 1394
FT TIT 1395
FT TIT 1396
FT TIT 1397
FT TIT 1398
FT TIT 1399
FT TIT 1400
FT TIT 1401
FT TIT 1402
FT TIT 1403
FT TIT 1404
FT TIT 1405
FT TIT 1406
FT TIT 1407
FT TIT 1408
FT TIT 1409
FT TIT 1410
FT TIT 1411
FT TIT 1412
FT TIT 1413
FT TIT 1414
FT TIT 1415
FT TIT 1416
FT TIT 1417
FT TIT 1418
FT TIT 1419
FT TIT 1420
FT TIT 1421
FT TIT 1422
FT TIT 1423
FT TIT 1424
FT TIT 1425
FT TIT 1426
FT TIT 1427
FT TIT 1428
FT TIT 1429
FT TIT 1430
FT TIT 1431
FT TIT 1432
FT TIT 1433
FT TIT 1434
FT TIT 1435
FT TIT 1436
FT TIT 1437
FT TIT 1438
FT TIT 1439
FT TIT 1440
FT TIT 1441
FT TIT 1442
FT TIT 1443
FT TIT 1444
FT TIT 1445
FT TIT 1446
FT TIT 1447
FT TIT 1448
FT TIT 1449
FT TIT 1450
FT TIT 1451
FT TIT 1452
FT TIT 1453
FT TIT 1454
FT TIT 1455
FT TIT 1456
FT TIT 1457
FT TIT 1458
FT TIT 1459
FT TIT 1460
FT TIT 1461
FT TIT 1462
FT TIT 1463
FT TIT 1464
FT TIT 1465
FT TIT 1466
FT TIT 1467
FT TIT 1468
FT TIT 1469
FT TIT 1470
FT TIT 1471
FT TIT 1472
FT TIT 1473
FT TIT 1474
FT TIT 1475
FT TIT 1476
FT TIT 1477
FT TIT 1478
FT TIT 1479
FT TIT 1480
FT TIT 1481
FT TIT 1482
FT TIT 1483
FT TIT 1484
FT TIT 1485
FT TIT 1486
FT TIT 1487
FT TIT 1488
FT TIT 1489
FT TIT 1490
FT TIT 1491
FT TIT 1492
FT TIT 1493
FT TIT 1494
FT TIT 1495
FT TIT 1496
FT TIT 1497
FT TIT 1498
FT TIT 1499
FT TIT 1500
FT TIT 1501
FT TIT 1502
FT TIT 1503
FT TIT 1504
FT TIT 1505
FT TIT 1506
FT TIT 1507
FT TIT 1508
FT TIT 1509
FT TIT 1510
FT TIT 1511
FT TIT 1512
FT TIT 1513
FT TIT 1514
FT TIT 1515
FT TIT 1516
FT TIT 1517
FT TIT 1518
FT TIT 1519
FT TIT 1520
FT TIT 1521
FT TIT 1522
FT TIT 1523
FT TIT 1524
FT TIT 1525
FT TIT 1526
FT TIT 1527
FT TIT 1528
FT TIT 1529
FT TIT 1530
FT TIT 1531
FT TIT 1532
FT TIT 1533
FT TIT 1534
FT TIT 1535
FT TIT 1536
FT TIT 1537
FT TIT 1538
FT TIT 1539
FT TIT 1540
FT TIT 1541
FT TIT 1542
FT TIT 1543
FT TIT 1544
FT TIT 1545
FT TIT 1546
FT TIT 1547
FT TIT 1548
FT TIT 1549
FT TIT 1550
FT TIT 1551
FT TIT 1552
FT TIT 1553
FT TIT 1554
FT TIT 1555
FT TIT 1556
FT TIT 1557
FT TIT 1558
FT TIT 1559
FT TIT 1560
FT TIT 1561
FT TIT 1562
FT TIT 1563
FT TIT 1564
FT TIT 1565
FT TIT 1566
FT TIT 1567
FT TIT 1568
FT TIT 1569
FT TIT 1570
FT TIT 1571
FT TIT 1572
FT TIT 1573
FT TIT 1574
FT TIT 1575
FT TIT 1576
FT TIT 1577
FT TIT 1578
FT TIT 1579
FT TIT 1580
FT TIT 1581
FT TIT 1582
FT TIT 1583
FT TIT 1584
FT TIT 1585
FT TIT 1586
FT TIT 1587
FT TIT 1588
FT TIT 1589
FT TIT 1590
FT TIT 1591
FT TIT 1592
FT TIT 1593
FT TIT 1594
FT TIT 1595
FT TIT 1596
FT TIT 1597
FT TIT 1598
FT TIT 1599
FT TIT 1600
FT TIT 1601
FT TIT 1602
FT TIT 1603
FT TIT 1604
FT TIT 1605
FT TIT 1606
FT TIT 1607
FT TIT 1608
FT TIT 1609
FT TIT 1610
FT TIT 1611
FT TIT 1612
FT TIT 1613
FT TIT 1614
FT TIT 1615
FT TIT 1616
FT TIT 1617
FT TIT 1618
FT TIT 1619
FT TIT 1620
FT TIT 1621
FT TIT 1622
FT TIT 1623
FT TIT 1624
FT TIT 1625
FT TIT 1626
FT TIT 1627
FT TIT 1628
FT TIT 1629
FT TIT 1630
FT TIT 1631
FT TIT 1632
FT TIT 1633
FT TIT 1634
FT TIT 1635
FT TIT 1636
FT TIT 1637
FT TIT 1638
FT TIT 1639
FT TIT 1640
FT TIT 1641
FT TIT 1642
FT TIT 1643
FT TIT 1644
FT TIT 1645
FT TIT 1646
FT TIT 1647
FT TIT 1648
FT TIT 1649
FT TIT 1650
FT TIT 1651
FT TIT 1652
FT TIT 1653
FT TIT 1654
FT TIT 1655
FT TIT 1656
FT TIT 1657
FT TIT 1658
FT TIT 1659
FT TIT 1660
FT TIT 1661
FT TIT 1662
FT TIT 1663
FT TIT 1664
FT TIT 1665
FT TIT 1666
FT TIT 1667
FT TIT 1668
FT TIT 16
```

```

RP SEQUENCE FROM N.A. (ALBIELE B).
RX MEDLINE=82173203; PubMed=6803173;
RA Olio R., Rougeon F.;
RT "Mouse immunoglobulin allotypes: post-duplication divergence of gamma
RL 2a and gamma 2b chain genes.";
RL Nature 296:761-763(1982).
[5]
RP CARBOHYDRATE-LINKAGE SITE THR-105.
RX MEDLINE=94216359; PubMed=7512967;
RA Kim H., Yamaguchi Y., Maeda K., Matsunaga C., Yamamoto K.,
RA Irimura T., Takahashi N., Kato K., Arata Y.;
RT "O-glycosylation in hinge region of mouse immunoglobulin G2b.";
RL J. Biol. Chem. 269:12345-12350(1994).
CC -1- SUBCELLULAR LOCATION: Secreted (potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoform=2;
CC Name=Secreted;
CC IsoId=P01866-1; Sequence=Displayed;
CC Note=May be the major isoform;
CC Name=Membrane-bound;
CC IsoId=P01867-1; Sequence=External;
CC MODIFIED WITH 2 BASIC ACID RESIDUES.
CC -1- PTM: DISULFIDE BONDS BY HOMOLOG WITH OTHER IG GAMMA CHAINS.
CC -1- MISCELLANEOUS: The a allele sequence is shown.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
DR HSBP, P01842; 7FAB.
DR InterPro: IPR007110; IG-1like.
DR InterPro: IPR003597; IG_c1.
DR InterPro: IPR003006; IG_MHC.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00407; IGc1_2.
DR PROSITE: PS50835; IG_LIKE; 3.
DR PROSITE: PS50290; IG_MHC; 1.
KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KM Alternative splicing; Repeat.
FT NON TER 1 1
FT DOMAIN 6 98 IG-LIKE 1.
FT 127 226 IG-LIKE 2.
FT DOMAIN 235 331 IG-LIKE 3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 118 118 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 150 210
FT DISULFID 256 314
FT CARBOHYD 105 105 O-LINKED (GALNAC...).
FT MOD RES 336 336 REMOVED POST-TRANSLATIONALLY (PROBABLE).
FT VARIANT 153 163 Q -> R (IN ALLELE B).
FT VARIANT 194 194 T -> A (IN ALLELE B).
FT VARIANT 300 300 N -> D (IN ALLELE B).
FT VARIANT 301 301 M -> I (IN ALLELE B).
FT VARIANT 301 301 L -> S (IN REF. 2 AND 3).
FT CONFLICT 25 25 S -> P (IN REF. 2 AND 3).
FT CONFLICT 36 36 I -> T (IN REF. 2 AND 3).
FT CONFLICT 239 239
SQ SEQUENCE 336 AA; 36658 MW; 7D879662607C356E CRC64;

Query Match 28.9%; Score 781.5; DB 1; Length 336;
Beet Local Similarity 49.8%; Pred. No. 5e-43; Indels 39; Gaps 9;
Matches 161; Conservative 50; Mismatches 73;

```

```

DB 133 PNTKDVLMISLTPKVCVVVDVSGDDVDVIGSFVNNVVEHTAQOTHRDYNSTIRVV 192
QY 289 SVLTVLHDDMLANGKEYKKVKNKALPAPIETKISAKQREPOYVTLPPSRDELTKQV 348
DB 193 STLPIQHDMMSGKEFKCKVNNKOLPSPBIERTISKIKLVAPQVYILPPAEQLSRKV 252
QY 349 SLTCLVKGFPSPDI AVENESNGOPENNYKTPPVLDSDGSRFLYSKLVVDSRMQGVF 408
DB 253 SLTCLVGFNPDLSVETTSNGHTBENKDYAPVLDSDGSFTISKLNKTKTSKWEKTSDF 312
QY 409 SCVWHEALHNHYTQKSLSLSPG 431
DB 313 SCVWHEALHNHYTQKSLSLSPG 335

RESULT 29
CD4_SAISC STANDARD; PRT; 457 AA.
AC 029037;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3).
GN CD4.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_TaxId=9521;
RN [1]
RP SEQUENCE FROM N.A.
RA Tachumi M., Hashimoto O.;
RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (by similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: D86588; BA013131.1; -.
DR HSBP, P01730; IMR.
DR GO: GO:0042101; C: T-cell receptor complex; ISS.
DR GO: GO:0015026; F: coreceptor activity; ISS.
DR GO: GO:0042289; F: MHC class II protein binding; ISS.
DR GO: GO:0006955; P: immune response; ISS.
DR GO: GO:0045086; P: positive regulation of interleukin-2 biosyn. .; ISS.
DR GO: GO:0030217; P: T-cell differentiation; ISS.
DR GO: GO:0045058; P: T-cell selection; ISS.
DR GO: GO:0007169; P: transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro: IPR000973; CD4_TcAg.
DR InterPro: IPR007110; IG_Like.
DR InterPro: IPR003596; IG_V.
DR Pfam: PF00047; Ig_2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IG_V.
DR PROSITE: PS50835; IG_LIKE; 1.
KM Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KM Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT DOMAIN 26 457 T-CELL SURFACE GLYCOPROTEIN CD4.
FT DOMAIN 26 395 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 396 417 POTENTIAL.
FT DOMAIN 418 457 CYTOPLASMIC (POTENTIAL).

```

```

FT DOMAIN 26 125 IG-LIKE V-TYPE.
FT DOMAIN 126 202 IG-LIKE C2-TYPE 1.
FT DOMAIN 316 203 IG-LIKE C2-TYPE 2.
FT DOMAIN 317 373 IG-LIKE C2-TYPE 3.
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 327 369 BY SIMILARITY.
FT LIPID 418 418 S-palmitoyl cysteine (By similarity).
FT LIPID 421 421 S-palmitoyl cysteine (By similarity).
SQ SEQUENCE 457 AA; 50871 MW; 57EED344005A015 CRC64;

```

```

Query Match 27.5%; Score 743; DB 1; Length 457;
Best Local Similarity 40.1%; Pred. No. 2.1e-40;
Matches 209; Conservative 56; Mismatches 150; Indels 106; Gaps 20;

```

```

OY 1 MNRGVPRHLLVLTQALLPAATQGNKVVLGKGGDTVELTCTASQKSIQFMKNSQIK 60
DB 1 MNGGIPFHHLLVLTQALLPAVTHGKTVVLGKGGVEVLPCTSLKKNVPEHMKTSQIK 60
OY 61 ILNGSGSLTQGPSTLNDNRADSRSLMDSQNFPLIKLTKEDSDTYICEVEDQKEVQL 120
DB 61 ILGVQNVFVYTGQSLTDRIDSKSSWDRGSPFLIKARLEDETYICEVESKKEVEL 120
OY 121 LVFGLTANSDTHLQGGSLTTLTLESPGSSPVQCSRPRKNIQGGKTLVSQLELDPSG 180
DB 121 QVFGTANPDTHLQGGSLTTLTLESPGSSPVQCSRPRKRIKGRKTLVSQLELDPSG 180
OY 181 TWCTVLTQNGKVEFKIDIVCPAPBPKSCDKHTCEPLLGGPSVFLFP-PRKQDTLMIS 239
DB 181 TMKCTVFOHLELV-FEINIVLAFQAS-----TVYKKEGEVPSFPLAFAETLTGS 234
OY 240 RTPETVCVVVDVSHDEPEVKNVYDGVENAKTKREEDYNTYRVSVLTYLHADM 299
DB 240 RTPETVCVVVDVSHDEPEVKNVYDGVENAKTKREEDYNTYRVSVLTYLHADM 299
OY 235 G-----ELCW-----QARRASSSKSWITFNLTKEVYKLT----- 266
DB 300 NGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGF 359
DB 267 --QPKLKMGRK--LPHLTLAQLPQYASSGNFTL-----ALKG--- 302
OY 360 SDIAVESWESNQEPENNYKTPPVLDSDGSFFLYSLKTVYDKSRMOQGVNFCGVMEALHN 419
DB 303 -----KT-----GKLHGEVNLVYMRVTQ--QNML--TEVWGP----- 332
OY 420 HYTKSLSTSPGLQUDERCARQDGEIDGWTTPR-PRASALPAPPGSALPDPQTASAL 478
DB 333 --TSPKMLLS--LKLNEQAKVSKRE-KAVVNLNPEPGAWQCLLSDSGQVLLSK-PEAL 386
OY 479 P-DPPAASALPALAVISFLGL-----GLGV-ACVLAATR 512
DB 387 PTRSPVQ---PWLIVIGVAGLLAFTGLGIFLCVRCRHR 424

```

```

RESULT 30
CD4 CANFA STANDARD; PRT; 463 AA.
AC P33705;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell) surface antigen
DE T4/Leu-3).
GN CD4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN [1]
RP SEQUENCE OF 13-463 FROM N.A.
RC STRAIN=Beagle; TISSUE=Thymus;
RX MEDLINE=93192324; PubMed=7916632;

```

```

RA Milde K.F., Conner G.E., Minz D.H., Alejandro R.;
RT "Primary structure of the canine CD4 antigen.";
RL Biochim. Biophys. Acta 1172:315-316(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Thymus;
RX MEDLINE=94378217; PubMed=8091416;
RA Gorman S.D., Frewin M.R., Cobbold S.P., Waldmann H.;
RT "Isolation and expression of cDNA encoding the canine CD4 and CD8
RT alpha antigens.";
RL Tissue Antigens 43:184-186(1994).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in macrophages and a subset of
CC T lymphocytes.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L06130; AB02295.1; -.
DR EMBL, X68565; -, NOT_ANNOTATED_CDS.
DR HSSP, P01730; IWR.
DR GO, GO:0042101; C:T-cell receptor complex; ISS.
DR GO, GO:0015026; F:coreceptor activity; ISS.
DR GO, GO:0042289; F:MHC class II protein binding; ISS.
DR GO, GO:0006955; P:immune response; ISS.
DR GO, GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO, GO:0030217; P:T-cell differentiation; ISS.
DR GO, GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro, IPR000973; CD4 TCRg.
DR InterPro, IPR007110; Ig-like.
DR InterPro, IPR003596; Ig_V.
DR Pfam, PF00047; Ig_3.
DR PRINTS, PR00692; CD4TCANTIGEN.
DR SMART, SM00406; IGV; 1.
DR PROSITE, PS50835; IG_LIKE_1.
DR Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KM Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 463 T-CELL SURFACE GLYCOPROTEIN CD4.
FT DOMAIN 25 401 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 402 423 POTENTIAL.
FT DOMAIN 424 463 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 124 IG-LIKE V-TYPE.
FT DOMAIN 125 211 IG-LIKE C2-TYPE 1.
FT DOMAIN 212 321 IG-LIKE C2-TYPE 2.
FT DOMAIN 322 378 IG-LIKE C2-TYPE 3.
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 332 374 BY SIMILARITY.
FT LIPID 424 424 S-palmitoyl cysteine (By similarity).
FT LIPID 427 427 S-palmitoyl cysteine (By similarity).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 389 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 463 AA; 51639 MW; 95805170CB44A833 CRC64;

```

```

Query Match 22.9%; Score 617.5; DB 1; Length 463;
Best Local Similarity 33.5%; Pred. No. 2.2e-32;
Matches 177; Conservative 68; Mismatches 155; Indels 129; Gaps 19;

```


QY 1 MNRGVPRFRLHLLVLTALPPAATGKNVYLGKGGDTVELTCTASQKSIQIHFMKNSNOIK 60
 DB 1 MNRGVPRFRLHLLVLTALPPAATGKNVYLGKGGDTVELTCTASQKSIQIHFMKNSNOIK 60
 QY 61 ILGNQSGFLTKPSPKLNDRADRSRLMDQGNFPLIINKLIDSDTYICEVEDQKEEVL 120
 DB 61 ILGNQSGFLTKPSPKLNDRADRSRLMDQGNFPLIINKLIDSDTYICEVEDQKEEVL 120
 QY 61 ILGNQSGFLTKPSPKLNDRADRSRLMDQGNFPLIINKLIDSDTYICEVEDQKEEVL 120
 DB 61 ILGNQSGFLTKPSPKLNDRADRSRLMDQGNFPLIINKLIDSDTYICEVEDQKEEVL 120
 QY 121 LVFGTLA-----NSDTHLLOGQSLLTLTSPGSSPVQCRSPGKNIQGGKTLIS 171
 DB 121 LVFGTLA-----NSDTHLLOGQSLLTLTSPGSSPVQCRSPGKNIQGGKTLIS 171
 QY 120 LVFNLTAKKDSGSSSSSSNIRLLQGGQLTLTENSGSSPVQMGPKNSHGQGNLSL 179
 DB 120 LVFNLTAKKDSGSSSSSSNIRLLQGGQLTLTENSGSSPVQMGPKNSHGQGNLSL 179
 QY 172 SOLLELDGSGTWTCTVLOQOKVYKFIIDVPCAPAPKSCDKHTCTPELIGPSVFLPFPK 231
 DB 172 SOLLELDGSGTWTCTVLOQOKVYKFIIDVPCAPAPKSCDKHTCTPELIGPSVFLPFPK 231
 QY 180 SMPBELDGGTWTCTIISQSKTVEFNINVLAF---QKSNFYARE--GDQVEISFP-- 232
 DB 180 SMPBELDGGTWTCTIISQSKTVEFNINVLAF---QKSNFYARE--GDQVEISFP-- 232
 QY 232 PVDITMISRTPEVTCVVVDVSHEDPEV--KFNWYVDGEVHNNAKTKPREQYNSTYRVVS 289
 DB 232 PVDITMISRTPEVTCVVVDVSHEDPEV--KFNWYVDGEVHNNAKTKPREQYNSTYRVVS 289
 QY 233 -----LSFEDENLVGELRMQAGAS-----SS 254
 DB 233 -----LSFEDENLVGELRMQAGAS-----SS 254
 QY 290 VITVLHODVLNGKEYKCKVSNKALPAPIEKTSKAKGPRE--POVYTLPSRDELTKNQ 347
 DB 290 VITVLHODVLNGKEYKCKVSNKALPAPIEKTSKAKGPRE--POVYTLPSRDELTKNQ 347
 QY 295 LL-----WISFTLENRKLSMKEAHAPL-----KLQKESLPURFTLPQVLSRYAGSG 301
 DB 295 LL-----WISFTLENRKLSMKEAHAPL-----KLQKESLPURFTLPQVLSRYAGSG 301
 QY 348 VSLTCLVKGFPYSDIAVENESNGQPENNYKTPPYLDSGSPFLYSKLTVDKSRMQGNV 407
 DB 348 VSLTCLVKGFPYSDIAVENESNGQPENNYKTPPYLDSGSPFLYSKLTVDKSRMQGNV 407
 QY 302 ILTLNLALGTLXQEV-----NLVWRANSQNNL 330
 DB 302 ILTLNLALGTLXQEV-----NLVWRANSQNNL 330
 QY 408 FSCSVMEALHNHYQKSLISLPGQLDETCAEADGDELGMWTTDPPASALPAPPTGS 467
 DB 408 FSCSVMEALHNHYQKSLISLPGQLDETCAEADGDELGMWTTDPPASALPAPPTGS 467
 QY 331 -TCEVLGP-----TSPELTIS--LNLKEQAQKVRK--QOKLVWVVDPEGCT-----WQC 374
 DB 331 -TCEVLGP-----TSPELTIS--LNLKEQAQKVRK--QOKLVWVVDPEGCT-----WQC 374
 QY 468 ALPDPQ---TASAL---PPPPASALPALA-----VISITLGLGIGVAC 506
 DB 468 ALPDPQ---TASAL---PPPPASALPALA-----VISITLGLGIGVAC 506
 QY 375 LLSDKDKVLLASLVSSPVVVKSWPKFLATLGLLIGLGLCVPC 423
 DB 375 LLSDKDKVLLASLVSSPVVVKSWPKFLATLGLLIGLGLCVPC 423

RESULT 31

CD4_RABIT STANDARD; PRT; 459 AA.

AC P46630;
 DT 01-NOV-1995 (rel. 32, Created)
 DT 01-NOV-1995 (rel. 32, Last sequence update)
 DT 10-OCT-2003 (rel. 42, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
 GN CD4.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_Taxid=9986;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92290370; PubMed=151821;
 RA Haque B.F., Sawadikosol S., Brown T.J., Lee K., Recker D.P., Kindt T.J.;
 RT "CD4 and its role in infection of rabbit cell lines by human immunodeficiency virus type 1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7963-7967(1992).
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-1ck (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----EMBL: M92840; AAA3198.1; --
 DR PIR: A46254; A46254.
 DR HSPB, P01730; 1MR.
 DR GO: GO:0042101; C:T-cell receptor complex; ISS.
 DR GO: GO:0015026; F:coreceptor activity; ISS.
 DR GO: GO:0042289; F:MHC class II protein binding; ISS.
 DR GO: GO:0006955; P:immune response; ISS.
 DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO: GO:0030217; P:T-cell differentiation; ISS.
 DR GO: GO:0045058; P:T-cell selection; ISS.
 DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro: IPR00973; CD4 TCRG.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003596; IG_v.
 DR Pfam: PF00047; 1g; 2.
 DR SMART; PRO0692; CD4TCANTIGEN.
 DR SMART; SMO0406; IGv; 1.
 DR PROSITE; PS0835; IG LIKE; 1.
 KM Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell; Immune response; Repeat; Signal; Lipoprotein; Palmitate.
 FT SIGNAL 1 25
 FT CHAIN 26 459
 FT DOMAIN 26 396
 FT TRANSMEM 397 419
 FT DOMAIN 420 459
 FT DOMAIN 26 129
 FT DOMAIN 130 208
 FT DOMAIN 209 318
 FT DOMAIN 319 374
 FT CAROHRD 299 299
 FT DISULFID 41 113
 FT DISULFID 329 370
 FT LIPID 420 420
 FT LIPID 423 423
 SQ SEQUENCE 459 AA; 50886 MW; B323311CBDA0013D CRC64;

Query Match 22.3%; Score 601.5; DB 1; Length 459;
 Best Local Similarity 41.9%; Pred. No. 2.3e-31;
 Matches 153; Conservative 47; Mismatches 98; Indels 67; Gaps 11;

QY 1 MNRGVPRFRLHLLVLTALPPAATGKNVYLGKGGDTVELTCTASQKSIQIHFMKNSNOIK 60
 DB 1 MNRGVPRFRLHLLVLTALPPAATGKNVYLGKGGDTVELTCTASQKSIQIHFMKNSNOIK 60
 QY 61 ILGNQSGFLTKPSPKLNDRADRSRLMDQGNFPLIINKLIDSDTYICEVEDQKEEVL 116
 DB 61 ILGNQSGFLTKPSPKLNDRADRSRLMDQGNFPLIINKLIDSDTYICEVEDQKEEVL 116
 QY 117 EVOLLVGLTNSDTHLLOGQSLLTLTSPGSSPVQCRSPGKNIQGGKTLISVQLEL 176
 DB 117 EVOLLVGLTNSDTHLLOGQSLLTLTSPGSSPVQCRSPGKNIQGGKTLISVQLEL 176
 QY 121 EVELLVFRLTANPTRLHAGSLTLTLEGPSVQVMSPEKNIETGPTCMPKRL 180
 DB 121 EVELLVFRLTANPTRLHAGSLTLTLEGPSVQVMSPEKNIETGPTCMPKRL 180
 QY 177 QDSGTWCTV-LONOKVYKFIIDVPCAPAPKSCDKHTCTPELIGPSVFLPFPKPD 235
 DB 177 QDSGTWCTV-LONOKVYKFIIDVPCAPAPKSCDKHTCTPELIGPSVFLPFPKPD 235
 QY 181 QDSGTWCTV-LONOKVYKFIIDVPCAPAPKSCDKHTCTPELIGPSVFLPFPKPD 229
 DB 181 QDSGTWCTV-LONOKVYKFIIDVPCAPAPKSCDKHTCTPELIGPSVFLPFPKPD 229
 QY 236 LMSRTPEVTCVVVDVSHEDPEV--KFNWYVDGEVHNNAKTKPREQYNSTYRVVS 293
 DB 236 LMSRTPEVTCVVVDVSHEDPEV--KFNWYVDGEVHNNAKTKPREQYNSTYRVVS 293
 QY 230 -----LNFEDSLGELMWQDGAS----- 249
 DB 230 -----LNFEDSLGELMWQDGAS----- 249
 QY 294 LHQWLNGKEYKCKVSNKALPAPIEKTSKAKGPREPOVYTLPSRDELTKNQSLT 351
 DB 294 LHQWLNGKEYKCKVSNKALPAPIEKTSKAKGPREPOVYTLPSRDELTKNQSLT 351
 QY 250 SAQSWSPSLDRKVSQKIP---DKIQMSKGLPLS---LTLPOALHRYAGSGNLSLT 303
 DB 250 SAQSWSPSLDRKVSQKIP---DKIQMSKGLPLS---LTLPOALHRYAGSGNLSLT 303
 QY 352 CLVKG 356
 DB 304 -LDKG 307

RESULT 32

CD4_RAT STANDARD; PRT; 457 AA.

AC P05540; 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3) (W3/25 antigen).
 GN CD4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_Taxid=10116;
 RX MEDLINE=87175535; PubMed=3104900;
 RA Clark S.J., Jeffries W.A., Barclay A.N., Gagnon J., Williams A.F.,
 RT "Peptide and nucleotide sequences of rat CD4 (W3/25) antigen:
 RT evidence for derivation from a structure with four
 RT immunoglobulin-related domains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:1649-1653(1987).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 210-393.
 RX MEDLINE=93262437; PubMed=8493535.
 RA Brady R.L., Dodson E.J., Dodson G.G., Lange G., Davis S.J.,
 RT "Crystal structure of domains 3 and 4 of rat CD4: relation to the
 RT NH2-terminal domain.";
 RL Science 260:979-983(1993).
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M15768; AAA0901.1; -
 DR PIR; A27449; A27449.
 DR PDB; 1CID; 15-JUL-93.
 DR GLCOSUNITEDB; P05540; -
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.
 DR GO; GO:0015026; F:coreceptor activity; ISS.
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.
 DR GO; GO:0006955; P:immune response; ISS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO; GO:0045058; P:T-cell differentiation; ISS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro; IPR000973; CD4-TCAG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; Ig; 2;
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00409; IG; 2;
 DR PROSITE; PS50835; IG-LIKE; 1.
 KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
 KW Immune response; Repeat; Signal; Lipoprotein; Palmitate; 3D-structure.
 FT SIGNAL 1 27
 FT CHAIN 28 457
 FT DOMAIN 28 394 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 395 417 POTENTIAL.
 FT DOMAIN 418 457 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 28 127 IG-LIKE V-TYPE.
 FT DOMAIN 128 206 IG-LIKE C2-TYPE 1.
 FT DOMAIN 207 316 IG-LIKE C2-TYPE 2.
 FT DOMAIN 317 374 IG-LIKE C2-TYPE 3.
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 43 111 BY SIMILARITY.
 FT DISULFID 158 187 BY SIMILARITY.
 FT DISULFID 328 370 BY SIMILARITY.
 FT LIPID 418 418 S-palmitoyl cysteine (By similarity).
 FT LIPID 421 421 S-palmitoyl cysteine (By similarity).
 FT STRAND 213 217
 FT TURN 218 219
 FT STRAND 222 225
 FT STRAND 235 243
 FT STRAND 252 258
 FT TURN 259 260
 FT STRAND 261 265
 FT STRAND 274 275
 FT STRAND 278 278
 FT TURN 279 280
 FT STRAND 282 285
 FT HELIX 290 292
 FT STRAND 294 301
 FT STRAND 306 319
 FT STRAND 325 331
 FT STRAND 338 344
 FT TURN 345 346
 FT STRAND 349 353
 FT STRAND 357 361
 FT STRAND 367 374
 FT TURN 375 376
 FT STRAND 377 385
 SQ SEQUENCE 457 AA; 51437 MW; 477BE157D30954C1 CRC64;
 Query Match 18.3%; Score 495; DB 1; Length 457;
 Best Local Similarity 32.8%; Pred. No. 1,5e-24;
 Matches 137; Conservative 57; Mismatches 132; Indels 92; Gaps 14;
 QY 1 MARGVPRRH--LLVQLALLPAAAGNKKVYLGKGDVLTCTASQKSIQPHKNSNQ 58
 DB 1 MCRGFSFRHLLPILLQLSKLVLTQGTVLGKGSALPEBSTSRASFAWKSSDQ 60
 QY 59 IKLNGSGSLTGPSTLNDRADSRSLMDQGNPLIKLKIEDSDTYCEVEDQEEV 118
 DB 61 KTLGVNKKLLIGSLSELYSPDSRKAMWGSPLINKLRMDSQTYVCELENKEEV 120
 QY 119 QLVFGLTANSDPHLLQGGSLTTLLES-PPGSSPSVOCSPRKNIOGGKTLISVSOLEQ 177
 DB 121 ELWVFRFTFNGRRLIQGGSLTILDSNPKVSPPIECKKSNIVDSKAFTHSLRIQ 180
 QY 178 DSGTWCTVLQNGKQVE--KIDIVPCAPDPKSCDTHTCPELLGSPVFLPPPKDPT 235
 DB 181 DSGIWNCTVLNOKKHSFDMKLSVL-----GFASTSIYAKSGESAEFSFP----- 227
 QY 236 LMTSRPEVTCVVVDVSHEDPEVKFMWYVNDGVEVNAKTKPRREQNSTRVIVLVLA 295
 DB 228 -----LNGESLQGELEW-----KAEKAPSS----- 249
 QY 296 QDWLNGKEYKCKVSNALPAPIEKTIKAKGQPREPOVYTLPPSRDELTRKNQVSLTCLVK 355
 DB 250 QSWITTSIKNQKYS-----VQKSTSNPKFQJSE---TLP-----LTIQI- 285
 QY 356 GYPSDIAVWESNQPENNYKTTTPVLDSDGSFPLYSKLTVDKSRWQGNVFSQVM 413
 DB 286 ----PQSIQFASG---NLTLT---LDR-GILYQEVNLVWVKVQPDSDNTLTCEVM 331
 RESULT 33
 CD4_MOUSE STANDARD; PRT; 457 AA.
 AC P06332;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3) (T-cell differentiation antigen L3T4).

GN CD4.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxId=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=87018845; PubMed=3094146;
 RA Touvaille B., Gorman S.D., Field E.H., Hunkapiller T., Parmes J.R.;
 RT "Isolation and sequence of L3T4 complementary DNA clones: expression
 in T cells and brain.";
 RL Science 234:610-614(1986).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=87115821; PubMed=3027575;
 RA Littman D.R., Gettner S.N.;
 RT "Unusual insertion in the immunoglobulin domain of the newly isolated
 murine CD4 (L3T4) gene.";
 RL Nature 325:453-455(1987).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX TISSUE=Brain;
 MEDLINE=88152875; PubMed=3326818;
 RA Parmes J.R., Hunkapiller T.;
 RT "L3T4 and the immunoglobulin gene superfamily: new relationships
 between the immune system and the nervous system.";
 RL Immunol. Rev. 100:109-127(1987).
 RN [4]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RX TISSUE=Brain;
 MEDLINE=88041159; PubMed=2823269;
 RA Gorman S.D., Touvaille B., Parmes J.R.;
 RT "Structure of the mouse gene encoding CD4 and an unusual transcript
 in brain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7644-7648(1987).
 RN [5]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=98112780; PubMed=9445485;
 RA Ansel-Lati M.A., Oeltjen J.C., Schwartz S., Zhang Z., Muzny D.M.,
 Lu J., Gorrell J.H., Chinnault A.C., Belmont J.W., Miller W.,
 Gibbs R.A.;
 RT "Comparative sequence analysis of a gene-rich cluster at human
 chromosome 12p13 and its syntenic region in mouse chromosome 6.";
 RL Genome Res. 8:29-40(1998).
 RN [6]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RX STRAIN=C57BL/6J; TISSUE=Mammary gland;
 MEDLINE=22388257; PubMed=12477932;
 RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Altschul S.F., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 Boeck S.A., McEwen P.J., McKernan K.O., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
 RN [7]
 RN SEQUENCE OF 27-43.
 RX MEDLINE=8616694; PubMed=3082751;
 RA Claesson B.J., Tesgaratos J., Kirsbaum L., Maddox J., McKay C.R.,
 Brandon M., McKenzie I.F.C., Walker I.D.;

RT "The L3T4 antigen in mouse and the sheep equivalent are
 RT immunoglobulin-like.";
 RL Immunogenetics 23:129-132(1986).
 RN [8]
 RN DISULFIDE BONDS.
 RX MEDLINE=86233454; PubMed=3086886;
 RA Claesson B.J., Tesgaratos J., McKenzie I.F.C., Walker I.D.;
 RT "Partial primary structure of the T4 antigens of mouse and sheep:
 RT assignment of intrachain disulfide bonds.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:4499-4503(1986).
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (by similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P06332-1; Sequence=Displayed;
 CC Name=2; Synonyms=Brain-specific;
 CC IsoId=P06332-2; Sequence=VSP 002489;
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; M36850; AAA39401.1; -;
 DR EMBL; M13816; AAA37267.1; -;
 DR EMBL; X04836; CAA28539.1; -;
 DR EMBL; M36851; AAA39402.1; -;
 DR EMBL; M17080; AAA37403.1; -;
 DR EMBL; M17078; AAA37403.1; JOINED.
 DR EMBL; M17079; AAA37403.1; JOINED.
 DR EMBL; AC002397; AAC36010.1; -;
 DR EMBL; BC039137; AAC39137.1; -;
 DR PIR; A02110; RMMST4.
 DR HSSP; P01730; IMBR.
 DR MGD; MGI:88335; Cd4.
 DR GO; GO:0042101; C: T-cell receptor complex; ISS.
 DR GO; GO:0015026; F: coreceptor activity; ISS.
 DR GO; GO:0042289; F: MHC class II protein binding; ISS.
 DR GO; GO:0006955; P: immune response; ISS.
 DR GO; GO:0045086; P: positive regulation of interleukin-2 biosyn. . .; ISS.
 DR GO; GO:0030217; P: T-cell differentiation; ISS.
 DR GO; GO:0045058; P: T-cell selection; ISS.
 DR GO; GO:007169; P: transmembrane receptor protein tyrosine kin. . .; ISS.
 DR InterPro; IPR0090973; CD4_TcAg.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 2.
 DR PRINTS; PR00692; CD4TCANTIGEN.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
 KW Immune response; Repeat; Signal; Lipoprotein; Palmitate;
 KW Alternative splicing.
 FT SIGNAL 1 26
 FT CHAIN 27 457
 FT DOMAIN 27 394
 FT TRANSMEM 395 417
 FT DOMAIN 418 457
 FT DOMAIN 27 128
 FT DOMAIN 129 207
 FT DOMAIN 208 317
 FT DOMAIN 318 374
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 374
 FT CARBOHYD 298 298
 FT CARBOHYD 323 323
 FT N-TERMINAL 1 26
 FT T-CELL SURFACE GLYCOPROTEIN CD4.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE V-TYPE.
 FT IG-LIKE C2-TYPE 1.
 FT IG-LIKE C2-TYPE 2.
 FT IG-LIKE C2-TYPE 3.
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (

FT	CARBOHYD	392	392	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	42	112	
FT	DISULFID	159	188	
FT	DISULFID	328	370	
FT	LIPID	418	418	S-palmitoyl cysteine (By similarity).
FT	LIPID	421	421	S-palmitoyl cysteine (By similarity).
FT	LIPID	421	421	Missing (in isoform 2).
FT	VARSPLIC	1	240	/FTID=VSP_002489.
SO	SEQUENCE	457 AA;	51296 MM;	1BD1A7527CB00F33 CRC64;
Query Match		17.6%;	Score 475;	DB 1; Length 457;
Best Local Similarity		52.8%;	Pred. No. 2.8e-23;	
Matches 104;	Conservative	32;	Mismatches 55;	Indels 6; Gaps 5;
QY	1	MNRGVPRRH-LLTVLQLALLPAATQGNKVYLCKGKDYELRTCTAQSQKSIQHMKNNOI	59	
Db	1	MCRAISLRLLLLQLSLQLLAVTQKTLVKGEESEALPESSQKQITVFTWTFSDOR	60	
QY	60	KILNQG-SFLTKG--PSKLNDRADSRSLMDQGNFPLIKLKIKEDSDTVICEVDQKE	116	
Db	61	KILQHGKGVILIRGSPSQF--DREPSKGNAMKESFPLINKLNKEDSDTVICELENKE	119	
QY	117	EVQLLVFGLTANSDTHLLQGSGLTITLES-PRGSSPVQCSRPRGNIKQGGKTLISVSOLE	175	
Db	120	EVELMVFKEVTPSPGTSILQGSGLTITLSDNSKVSNPLETECKHKKGVVSGSVLSMSNLR	179	
QY	176	LQDSGTWCTVTLQNOCK	192	
Db	180	WQSDDFMNCVTLLDQCK	196	

RESULT 34	
EPC_MOUSE	
ID_ EPC_MOUSE	STANDARD;
	PRT; 421 AA

```

DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 19 epsilon chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84236092; Pubmed=6329728;
RA Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;
RT "The nucleotide sequence of the mouse immunoglobulin epsilon gene:
RT comparison with the human epsilon gene sequence.";
EMBO J. 1:1117-1123(1982).
[2]
RP REVISIONS.
RA Honjo T.;
RL Submitted (APR-1986) to the EMBL/Genbank/DBJ databases.
[3]
RP SEQUENCE OF 34-421 FROM N.A.
RX MEDLINE=83117774; Pubmed=6618553;
RA Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
RT chain cDNA.";
Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
DR EMBL; X01857; CAA25977.1; -.
DR EMBL; X01857; CAA25978.1; -.
DR PIR; A02144; EHMS.

```

DR	PIR: A02145; EHMSS.
DR	HSSP; P01854; 1ICE.
DR	InterPro; IPR007110; Ig-like.
DR	InterPro; IPR003597; Ig_c1.
DR	InterPro; IPR003006; Ig_MHC.
DR	Pfam; PF00047; Ig1; 4.
DR	SMART; SM00407; Ig1; 2.
DR	PROSITE; PSS0835; IG_Like; 4.
DR	PROSITE; PS00230; IG_MHC; 3.
KW	Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT	NON TER
FT	DOMAIN 1 90
FT	DOMAIN 91 197
FT	DOMAIN 198 304
FT	DOMAIN 305 421
FT	DISULFID 23 75
FT	DISULFID 121 180
FT	DISULFID 226 285
FT	DISULFID 330 392
FT	CARBOHYD 43 43
FT	CARBOHYD 72 72
FT	CARBOHYD 84 84
FT	CARBOHYD 95 95
FT	CARBOHYD 166 166
FT	CARBOHYD 238 238
FT	CARBOHYD 261 261
FT	CARBOHYD 365 365
FT	CARBOHYD 415 415
FT	SEQUENCE 421 AA; 47320 MW; 8F909EIF3DA06B47 CRC64;

Query Match	14.4%;	Score 388;	DB 1;	Length 421;
Best Local Similarity	28.3%;	Pred. No. 9e-18;		
Matches 119;	Conservative 77;	Mismatches 138;	Indels 86;	Gaps 20

QY	38	ELTCTASOKXSIOGHMKNOSNDIKILNQGSLTGTGPGSLNDRADRSRLBMOQGNPLIIX	97
		55 ELKVTTSQVTS-----WKSAR-----NFTCHVTHPPSPNESRT-----ILVR	92
QY	98	NKIEDSDFTY-----CEVEDQKEEVOL--LVFGILTAN--SDTHLLOOSLTLTLESPPGS	149
Db	93	PVNITEPTLELHSSCDPNMFHSTIOLYCFYGHILNDVSYWMLDREIDTTL-----	146
QY	150	SPSVQCSPPGKNIOGKTLVSQLELQDSGTMTCTYLOKQKVEFKIDIVCPAPERKS	209
Db	147	AQVLVKE--EGKLASTGSKNTITEQOMWSESTPCKV--TSQGDYULNHTRCPDHEBR-	202
QY	210	CDKHTCEPELLGGSVFLPRPKMDTLMIKSTPEVTCVVDV--SHDEPVKN-----	261
Db	203	-----GYTTLIPSPD--LYQNGAPKLTCLVDLESEKNVNVYNNEKTSV	249
QY	262	-----WYVDGVEVHAKTKPREEGYSTYRVVSVLTVLHODMLANGEXKCKSNKALPPI	317
Db	250	SASQWY---TKHNN-----NATTSITSLIPVAKXMIEGYQCIYDHPRPPI	296
QY	318	EKTSKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAMWESNGOPENN--	375
Db	297	VRSITKTPGQASAPAEVYVPPPEER--SEDKRTTLCLIONFPEDISVQMLDGKLISSQ	355
QY	376	YKTPPLVDSDG---FFLYSKLTVDKSRMOQGVSCSVNHEALHN--HYQKSLSLSPG	431
	356	HSITTP--LKNGSNGQFFISRLLEVATLWQRIQFCOVYHEALQKRXKLEKTIISTLG	414

RESULT 35			
ID	MUC_HUMAN	STANDARD;	PRT; 454 AA.
AC	P01877;		
DT	21-JUL-1986	(Rel. 01, Created)	
DT	01-FEB-1991	(Rel. 17, Last sequence update)	
DT	15-MAR-2004	(Rel. 43, Last annotation update)	
DE	Ig mu chain C region.		
GN	IGHM.		
OS	Homo sapiens (Human).		

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-434 FROM N.A.
 RX MEDLINE=90332450; PubMed=2115996;
 RA Friedlander R.M., Nussenzweig M.C., Leder P.;
 RT "Complete nucleotide sequence of the membrane form of the human Igm
 RT heavy chain.";
 RL Nucleic Acids Res. 18:4278-4278(1990).
 RN [2]
 RP SEQUENCE (WALDENSTROM'S MACROGLOBULIN GAL).
 RX MEDLINE=75059123; PubMed=4803843;
 RA Watanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;
 RT "The primary structure of a monoclonal Igm-immunoglobulin
 RT (macroglobulin Gal.). II: the amino acid sequence of the H-chain (mu-
 RT type), subgroup H III. Architecture of the complete Igm-molecule.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).
 RN [3]
 RP REVISIONS (GAL).
 RX MEDLINE=81066716; PubMed=6777162;
 RA Mhaesco E., Barnikol-Watanabe S., Barnikol H.U., Mhaesco C.,
 RA Hilschmann N.;
 RT "The primary structure of the constant part of mu-chain-disease
 RT protein BOT.";
 RL Eur. J. Biochem. 111:275-286(1980).
 RN [4]
 RP SEQUENCE (WALDENSTROM'S OU), DISULFIDE BONDS, AND CARBOHYDRATES.
 RX MEDLINE=74005511; PubMed=4742735;
 RA Putnam F.W., Florent G., Paul C., Shinoda T., Shimizu A.;
 RT "Complete amino acid sequence of the Mu heavy chain of a human Igm
 RT immunoglobulin.";
 RL Science 182:287-291(1973).
 RN [5]
 RP PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE=82059479; PubMed=6795593;
 RA Rabbitts T.H., Forster A., Milstein C.P.;
 RT "Human immunoglobulin heavy chain genes: evolutionary comparisons of
 RT C mu, C delta and C gamma genes and associated switch sequences.";
 RL Nucleic Acids Res. 9:4509-4524(1981).
 RN [6]
 RP SEQUENCE OF 299-387 AND 438-454 FROM N.A.
 RX MEDLINE=81077306; PubMed=6777778;
 RA Dolby T.W., Devuono J., Croce C.M.;
 RT "Cloning and partial nucleotide sequence of human immunoglobulin mu
 RT chain cDNA from B cells and mouse-human hybridomas.";
 RL Proc. Natl. Acad. Sci. U.S.A. 77:6027-6031(1980).
 CC -1- MISCELLANEOUS: All 4 combinations of the S/G and V/G polymorphisms
 CC at positions 192 and 216 have been observed in human mu chains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, X17115; CAA34971.1; ALT SEQ.
 DR EMBL; X57086; -; NOT_ANNOTATED_CDS.
 DR HSSP; P01857; 1FC1.
 DR Genew; HGNC:5541; IGHN.
 DR MIM; 147020; -;
 DR GlycositeDB; P01871; -;
 DR GO; GO:0005624; C:membrane fraction; NAS.
 DR GO; GO:0003823; P:antigen binding; TAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-1like.
 DR InterPro; IPR003587; Ig_C1.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig_4.
 DR SMART; SM00407; IgC1; 3.
 DR PROSITE; PS00835; IG_LIKE; 4.

DR PROSITE; PS00290; IG_MHC; 3.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW Polymorphism.
 FT NON_TER 1 1
 FT DOMAIN 1 105 CH1.
 FT DOMAIN 106 218 CH2.
 FT DOMAIN 219 324 CH3.
 FT DOMAIN 325 454 CH4.
 FT DISULFID 14 14 INTERCHAIN (WITH LIGHT CHAIN).
 FT DISULFID 28 88
 FT DISULFID 135 198
 FT DISULFID 215 215 INTERCHAIN (WITH HEAVY CHAIN).
 FT DISULFID 245 304 INTERCHAIN (WITH HEAVY CHAIN IN ANOTHER
 FT DISULFID 292 292 SUBUNIT).
 FT DISULFID 352 414 INTERCHAIN (WITH HEAVY CHAIN).
 FT DISULFID 453 453 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .).
 FT CARBOHYD /FTid=CAR_000219.
 FT VARIANT 192 192 S -> G.
 FT VARIANT /FTid=VAR_003903.
 FT VARIANT 216 216 V -> G (in dbSNP:12365).
 FT VARIANT /FTid=VAR_003904.
 FT VARIANT 216 216 /FTid=VAR_003904.
 SQ SEQUENCE 454 AA; 49556 MW; 21EC72BAD056922E CRC64;
 Query Match 14.2%; Score 384.5; DB 1; Length 454;
 Best Local Similarity 24.9%; Pred. No. 1.7e-17;
 Matches 114; Conservative 73; Mismatches 164; Indels 107; Gaps 15;
 QY 36 TWELCTLSQ--KSIOTFMKNSNOIKILGNQSLTTPSKLNDRAISRSLMDQNF 93
 DB 23 SVAAGCLQDFLPDITFSWKYKNSDISSTRG-----FP 57
 QY 94 LIHKLKIEDSDTYCEVED-----QKEEYQLVFLGLTANSDDLHQQGSLTLTLESPPS 149
 DB 58 SVLRGKTAATISQVILPISKDMQGTDEHVCKVQHPNGKKNV---PLPIALHPKV 113
 QY 150 SPVOCR-----SPR-----GKNIQGGKTLVSQLELDQSG 180
 DB 114 SVFVPRPGFFGNPRSKSLIQATGFSRQVSWLABGKQVGSVTTDQQAIAKESG 173
 QY 181 -----TWCTVLQNGKVEFKIDVPCPAPEPKSCDKHTCPPEL 219
 DB 174 PTTVKTSTLTATKESDMLSQSWFTCRV--DHRGLTFQONASMCVDDPTAIR----- 224
 QY 220 LGPSVFLPFPKPKDTLMISRPETVTVVSHDEPKFPMVYDGVENNAKTKPRE 279
 DB 225 ----VFALPES-FASIFLTSTKLTCLVTLDTTYD-SVTLISWTRNGSAVKTHTNIS 277
 QY 280 QYNSTRVSVTLTVADHMLNKEKCYKSNALPAIEKTIKAKGP-REPOVYTLPP 338
 DB 278 HENAFSAVGEISIEDDMNSGERFTCYVTHDDLPPLAKQTSRKGVLAHRPDVYLLP 337
 QY 339 SHDELTKNOVSLTGLVGFPSDIAVWESNGQP-ENNYKTTTPVID--SDGSFFLYS 393
 DB 338 AREQNLRESATITCLTVGFSPADVFQVMQGGQLSPPEKYVTSAPMEPQAPGRYFMS 397
 QY 394 KLTVDKSRMOGANVSCSVMEHALNNHTQSLSLSPG 431
 DB 398 ILTVSEEMWTGETYTCVAHAHALPNRTERTVDSKG 435
 RESULT 36
 MISC_MOUSE
 ID MISC_MOUSE STANDARD; PRT; 476 AA.
 AC P01873;
 DT 21-JUL-1986 (rel. 01, Created)
 DT 01-NOV-1991 (rel. 20, Last sequence update)
 DT 10-OCT-2003 (rel. 42, Last annotation update)

```

OS Ig mu chain C region membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE OF 433-476 FROM N.A.
RX MEDLINE=80222874; PubMed=6771020;
RA Early P., Rogers J., Davis M., Calame K., Bond M., Wall R.,
RA Hood L.;
RT "Two mRNAs can be produced from a single immunoglobulin mu gene by
RT alternative RNA processing pathways.";
RL Cell 20:313-319(1980).
RN [2]
RP SEQUENCE OF 410-476 FROM N.A. (MPELONA MOPC 104E).
RX MEDLINE=80222873; PubMed=6771019;
RA Rogers J., Early P., Carter C., Calame K., Bond M., Hood L.,
RA Wall R.;
RT "Two mRNAs with different 3' ends encode membrane-bound and secreted
RT forms of immunoglobulin mu chain.";
RL Cell 20:303-312(1980).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=During differentiation, B lymphocytes switch from
CC expression of isoform Membrane-bound to isoform Secreted;
CC Name=Membrane-bound;
CC IsoId=P01873-1; Sequence=Displayed;
CC Name=Secreted;
CC IsoId=P01872-1; Sequence=External;
CC -1- MISCELLANEOUS: The sequence of residues 1-409 is assumed to be
CC identical with the corresponding region of the secreted form.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V00821; CAA24202.1; -.
CC PIR; A02167; MHMSM.
DR HSP, P01857, 1FCL.
DR InterPro; IPR007110, Ig-like.
DR InterPro; IPR003597, Ig_c1.
DR InterPro; IPR003006, Ig_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IGL1; 2.
DR PROSITE; PS50835; IG-LIKE, 4.
DR PROSITE; PS02900, IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT FT 1 105 CH1.
FT DOMAIN 106 217 CH2.
FT DOMAIN 218 324 CH3.
FT DOMAIN 325 436 CH4.
FT TRANSMEM 456 473 POTENTIAL.
FT DISULFID 14 14 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFID 28 89 BY SIMILARITY.
FT DISULFID 136 199 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 216 216 BY SIMILARITY.
FT DISULFID 246 305 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 293 293 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 353 415 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .).
FT CARBOHYD 281 281 N-LINKED (GLCNAC. . .).
SQ SEQUENCE 476 AA; 52656 MW; 8D476575A5204071 CXC64;

Query Match 13.8%; Score 373.5; DB 1; Length 476;
Beet Local Similarity 25.2%; Pred. No. 8.9e-17;

```

Match	121:	Conservative	87:	Mismatches	157:	Indels	115:	Gaps	23
QY	37	VELTCTASQ--KKSIIQFHWKNSNOIKILGNQGSFLTYGPKSLNDRADSRSLMDQNF--	92						
Db	24	VAMGCLARDELFPSTISFTWNYQNNTVEYLCIGIRTEPT-----LRTGCKYLA	68						
QY	93	---PLIIKNLKIIESDPNY-ICEVDDQKEBQVLLVGLTANSDFHL---LOGQSILTLLE	144						
Db	69	TSQVLLSPKSLIESGSDLEYLVCKIH-----YG-GKNRDLHVPIPAVAKMNPVNVF	117						
QY	145	SP-----GSSP-----SVOCR-----SPR-----GKNIQGG-----	166						
Db	118	VPPRDSGSGPARPRKSKILCEATNTPTKPIRTYSMLKDGVLVESGTTDPVTIENKSGTPTQ	177						
QY	167	---KTLVSQLELQDSGWTCTVLQNQKVEKIDIVCPAPRPKSCDKTHTEPELLGG	222						
Db	178	YKVIATLTLSIIDIMLNINVTYCRV--DHRGLTFLKVNYSSTCAASPT-----DIL-	225						
QY	223	PSVFLPPPKPMDTLMISRTPEVTCVWVDVSHDEPVKFNWVVDGVEVNNATKREEDYN	282						
Db	226	--TTTIPSPFAD-ITFLSKNALITCLVSNLAITYE-ITLNI SVMASOGSEPLETKIKIMESH	281						
QY	283	STYVSVSLVTLTLDHMDNGKYEKKCVSNKNKLPAPIENTISAKAQOPRE-----POVYTL	337						
Db	282	GTFSAKGVASVCEVDMNNRKEFTVTYHRDLPSPQKAFISK---PNEVHGNPRAVYILP	337						
QY	338	PSRDELTP-KNOVSLTLCVKGTFYPSDIAVEMESNQ--PENNYKTPTPVLD--SDGSFPLY	392						
Db	338	PABQOLNLRBSAATVTLCLVKGSPFADISVQWLQRGQLLPQEKYVTSAPMPERGAGAFETH	397						
QY	393	SKLVYDKSRNQGVNFGSCSVNHEMLNHNHYOKSLSLSPGLQDLDTCAEADQGEIDGLWTT	452						
Db	398	SILVTVTEBENVNSGETYTCVGVGHEALPHLYVERTYDKS---TEGEVNAABEGGFENLWTT	452						

RESULT 37

EPC_RAT

ID	EPC_RAT	STANDARD;	PRT;	429	AA.
AC	P01855;				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	19 epsilon chain C region.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RE	SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).				
RP	STRAIN=LOU/C/MSL;				
RC	MEDLINE=83064537;	Pubmed=6292865;			
RA	Heilmann L., Petersson U., Engstrom A., Karlsson T., Bennich H.;				
RT	"Structure and evolution of the heavy chain from rat immunoglobulin				
RL	E.";				
RL	Nucleic Acids Res. 10:6041-6049(1982).				
RN	[2]				
RP	SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).				
RX	MEDLINE=83182019;	Pubmed=6820340;			
RA	Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;				
RT	"A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:				
RL	construction, identification, and DNA sequence."				
RL	DNA 1:335-343(1982).				
RN	[3]				
RP	SEQUENCE OF 205-306 FROM N.A.				
RX	MEDLINE=82174576;	Pubmed=6802338;			
RA	Heilmann L., Petersson U., Bennich H.;				
RT	"Characterization and molecular cloning of the mRNA for the heavy				
RL	(epsilon) chain of rat immunoglobulin E.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982).				
CC	-I-SIMILARITY: Contains 4 immunoglobulin-like domains.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				


```

Db 251 VTDLITVYG-SLINISW-----ASHNGKALDTHMNTESHSPNATFSAMGEASVCAEDMESGE 304
Qy 303 EYKCVSKKALPAPLEKTIISAKGPRE-POVYTLPPSRDEL-T-KNOVSLTCLVKGFFPS 360
Db 305 QPTCVTHADLPFPFKHTISKSREVAKHPPAYVLPAPREQLVLESATVTCVLVGFSPA 364
Qy 361 DIAVESNNGOP--ENNKTTPPVLD--SDGSFFLYSKLTVDKSRMOQGNVSCSVMEBA 416
Db 365 DVFVQMGQRGQPLSDKVTISAPAPEPQAPGLYFTHTSLTYTEEDMNGGEFTTCVVGHEA 424
Qy 417 LHNHYTKQSLSLSPGLQDFTCAEPQDGLDGLMTT 452
Db 425 LPHWYERTVDS-----TEGEVGAEEGFEMLMTT 455

```

RESULT 39

```

MUC_CANFA STANDARD; PRT; 450 AA.
ID MUC_CANFA
AC P01874;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig mu chain C region.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE OF 1-177 (MOO).
RX MEDLINE=80077682; PubMed=117299;
RA McCumber L.J., Capra J.D.;
RT "The complete amino-acid sequence of a canine mu chain.";
RL Mol. Immunol. 16:565-570 (1979).
RN [2]
RP SEQUENCE OF 178-450 (MOO).
RX MEDLINE=78180587; PubMed=653360;
RA Wasserman R.L., Capra J.D.;
RT "Amino acid sequence of the Fc region of a canine immunoglobulin M: interspecies homology for the Igm class.";
RL Science 200:1159-1161 (1978).
DR PIR; A83131; MHDG.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR007110; IG-1like.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IgC1; 2.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON TER
FT SEQUENCE 450 AA; 48895 MW; 9DA60DA9D012F5D CRC64;

```

```

Query Match 13.7%; Score 369; DB 1; Length 450;
Best Local Similarity 24.1%; Pred. No. 1.6e-16;
Matches 119; Conservative 80; Mismatches 139; Indels 166; Gaps 20;

Qy 36 TVELTCTASQ--KSGIOFHMKSNQIKILGNQGSTLTGPKLNDRAISRSLMDQGNFP 93
Db 22 TVAMGCLARDPLPGSITFSMKYBLSTAINSTRG-----FP 56
Qy 94 LIILKLK-----IEDSDTYI-CEVE-----DOKEVQLVFGLTANSPTHLL 134
Db 57 SVLRGKGVANSQVLPBVDIIGTDEHIVCKVRHSBKKQKVPVPM----- 105
Qy 135 QGQSITLTLESPPGSSPSVQCR-----SPR-----GKNIOGSKT 168
Db 106 -----LTL--PPEVSGFIPRDAFFGPRKSQLICQASGSPRQVMSLRQKQIESGVT 157
Qy 169 LSVSGLLEQDSG-----TWTCT-----TVLONQKVEFKDIV 200
Db 158 TNEVAZAKZSGPPTTKVTSMLTIOEDAMLSQSVFTCVHRGLTFQONASM----- 210

```

```

Qy 201 PCPAPBPKSCDKTHTCEPLGSPSVFLPPPKKDTLMISRTPEVTVVDVSHEDPEYK 260
Db 211 -CTSDQGV-----GISIFITLPPS-FASINTNTSAKSLCVTLDTLATD-SVTI 254
Qy 261 NWYDVGEVHNATKPREEQNSTYRVVSLTVLHQDLNGKEYCKKCVSKALPAPLEKT 320
Db 255 SWTEENGALKTHTNISESHPNNGTFPSAMGEATVCEEMESGEQPTCVTHDLPVSKOT 314
Qy 321 ISXKKG-QPREPOVYTLPPSRDEL-TKNOVSLTCLVKGFFPSDIAVESNNGO--PENNY 376
Db 315 ISRPKGVAVHNPVYVLPSPREQLDLRESATLSCLTVGYSPPDVQVQVQKQVPPVPSY 374
Qy 377 KTTTPPVLD--SDGSFFLYSKLTVDKSRMOQGNVSCSVMEBALNHYTKQSLSPG--- 431
Db 375 VTSAPMEPQAPGLYFHTSLTYTEEDMNGGEFTTCVVAHESLPRNRYERSVDSTGKPT 434
Qy 432 -----IQDFTCAE 440
Db 435 LYNVSLVSDTAGZ 448

```

RESULT 40

```

MUCB_HUMAN STANDARD; PRT; 391 AA.
ID MUCB_HUMAN
AC P04220;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig mu heavy chain disease protein (BOT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=84184186; PubMed=6425189;
RA Barnikol-Watanabe S., Mihaesco E., Mihaesco C., Barnikol H.U.,
RA Hilschmann N.;
RT "The primary structure of mu-chain-disease protein BOT. Peculiar amino-acid sequence of the N-terminal 42 positions.";
RL Hoppe-Seyler's Z. Physiol. Chem. 365:105-118 (1984).
CC -I- MISCELLANEOUS: This protein has no V region homology or CH1 region.
DR PIR; A02163; MHHUBT.
DR HSSP; P01857; 1FC1.
DR GO; GO:0005624; C:membrane fraction; NAS.
DR GO; GO:0003823; F:antigen binding; TAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; IG-1like.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IgC1; 2.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region.
FT DOMAIN 1 42 PRE-C-PART (NO V REGION HOMOLGY).
FT DOMAIN 43 155 CH2.
FT DOMAIN 156 261 CH3.
FT DOMAIN 262 391 CH4.
FT SEQUENCE 391 AA; 43057 MW; 9100843AF0CF021A CRC64;

```

```

Query Match 13.5%; Score 365.5; DB 1; Length 391;
Best Local Similarity 28.0%; Pred. No. 2.3e-16;
Matches 105; Conservative 65; Mismatches 132; Indels 73; Gaps 15;

Qy 110 EVED-----QKEEVQLVFGLTANSPTHLLTLESPPGSSPSVQCR----- 156
Db 18 EAEDRIITKEEARL-----SGRD--MOUTSQGVIALPFPKVSVPVPRRGFGFGRPKS 68
Qy 157 -----SPR-----GKNIOGSKTLSVQLELQDSG-----TWTCTVLQN--- 189
Db 69 KLICQATGFSPPQLEVSWLRGKQVGSVTTTDEVEAAKSSGPTTYKVTSTLTIKESDWL 128

```



```
OY 190 -OKVEFKIDIVPCPAPEKSCDKHTTCPELLGSPS-----VLEPPKPKDTLMISRT 242
DB 129 GGSMTFCRVDHGLTQGNASS-----MCGEPQDRAIVFAIPSP-FASIFLTXT 178
OY 243 EVTCVVVDVSHEDPEVKFNNVYDGEVFNAAKTPKEEQYNSYRVVSVLTVLDHOMLNCK 302
DB 179 KLTCLTVLDTLYVD-SVTLSMTRODGEAVKHTHINISSEHNATFSAVGEASICEDDWDSGE 237
OY 303 EKCKVSNKALPAPIEKITSKAKGP-REPOVYTLTPSPDELT-KQVSLTCLVKGFPYS 360
DB 238 RETCTVHTHDLPSPLKQITISRPKGVALHRPDVLLPPEAQNLRESATITCLVYGFSPA 297
OY 361 DIAVEMESNGOP--ENNYKTTTPPLVD--SDGSFPLSKLTVDKSRMQGNVPSGSVMHEA 416
DB 298 DVFQVMQMGKGPLSPKVTYTSAPMEPPQAPGRYFAHSILTVEEEMTGETTCTVVAHEA 357
OY 417 LHNHYTQKSLSLSPG 431
DB 358 LFNRYVTERTVDSKG 372
RESULT 41
MUC_MOUSE STANDARD; PRT; 455 AA.
ID MUC_MOUSE
AC 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DB Ig mu chain C region secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCB_TaxId=10090;
(1)
RP SEQUENCE FROM N.A.
RX MEDLINE=81076590; PubMed=6255422;
RA Kawakami T., Takahashi N., Honjo T.;
RT "Complete nucleotide sequence of mouse immunoglobulin mu gene and
comparison with other immunoglobulin heavy chain genes.";
RL Nucleic Acids Res. 8:3931-3945(1980).
(2)
RP SEQUENCE FROM N.A.
RX MEDLINE=82051295; PubMed=6795090;
RA Goldberg G.I., Varin E.F., Zrolka A.M., Blattner F.R.;
RT "Sequence of the gene for the constant region of the mu chain of
Balb/c mouse immunoglobulin.";
RL Gene 15:33-42(1981).
(3)
RP SEQUENCE FROM N.A. (MYELOMA TEPCL83).
RX MEDLINE=81165562; PubMed=6260591;
RA Auftray C., Rougeon F.;
RT "Nucleotide sequence of a cloned cDNA corresponding to secreted mu
chain of mouse immunoglobulin.";
RL Gene 12:77-86(1980).
(4)
RP SEQUENCE (MYELOMA PROTEIN MOPC 104E).
RX MEDLINE=79233904; PubMed=111247;
RA Kehry M.R., Sibley C.H., Fuhrman J.S., Schilling J.W., Hood L.E.;
RT "Amino acid sequence of a mouse immunoglobulin mu chain.";
Proc. Natl. Acad. Sci. U.S.A. 76:2932-2936(1979).
(5)
RP REVISION (MOPC 104E).
RX MEDLINE=83075344; PubMed=6816276;
RA Kehry M.R., Fuhrman J.S., Schilling J.W., Rogers J., Sibley C.H.,
Hood L.E.;
RT "Complete amino acid sequence of a mouse mu chain: homology among
heavy chain constant region domains.";
Biochemistry 21:5415-5424(1982).
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Comment=During differentiation, B lymphocytes switch from
```

```
CC expression of isoform Membrane-bound to isoform Secreted;
CC Name=Secreted;
CC IsoId=P01872-1; Sequence=Displayed;
CC Name=Membrane-bound;
CC IsoId=P01873-1; Sequence=External;
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: V00827; -; NOT_ANNOTATED_CDS.
DR PIR: A02166; MHMS.
DR HSP: P01857; 1FCL.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_4.
DR SMART: SM00407; IG_C1_2.
DR PROSITE: PS50835; IG-LIKE; 4.
DR PROSITE: PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 105 CH1.
FT DOMAIN 106 217 CH2.
FT DOMAIN 218 324 CH3.
FT DOMAIN 325 436 CH4.
FT DOMAIN 437 455 C-TERMINAL SEGMENT OF SECRETED FORM.
FT DISULFID 14 14 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFID 28 89 BY SIMILARITY.
FT DISULFID 136 199 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 216 216 BY SIMILARITY.
FT DISULFID 246 305 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 293 293 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 353 415 BY SIMILARITY.
FT DISULFID 454 454 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .).
FT CARBOHYD 281 281 N-LINKED (GLCNAC. . .).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .).
FT VARIANT 78 78 S -> N (IN MOPC 104E).
FT VARIANT 101 101 H -> Q (IN MOPC 104E).
FT VARIANT 226 226 T -> N (IN TEPCL83 AND MOPC 104E).
FT VARIANT 258 258 N -> T (IN MOPC 104E).
FT VARIANT 258 258 N -> S (IN TEPCL83).
FT VARIANT 368 368 L -> K (IN TEPCL83 AND MOPC 104E).
SQ SEQUENCE 455 AA; 50101 MW; 4CB57CBB02F9B51 CRC64;
Query Match 13.4%; Score 362; DB 1; Length 455;
Best Local Similarity 25.3%; Pred. No. 4,6e-16;
Matches 116; Conservative 84; Mismatches 149; Indels 110; Gaps 22;
OY 37 VELTCTASQ--KKSIOFHKNNSQIKILGNQGSFLTKPSKLNDRADRRSLMDQGNF-- 92
DB 24 VAMGLADFLPSTISFTWNYQNTVEVIGRTFT-----LRTGKYLA 68
OY 93 ---PLIKLKIKEDDTY-ICEVEDQKEVQLLVGLTANSPTHL-----LQQSILTLTLE 144
DB 69 TSQVLLSPKSLIEGDEYLVCKH-----YG-SKNRDLHVPDPAVAMNPNVNF 117
OY 145 SPP-----GSSP-----SVOCR-----SPR-----GNKIQG----- 166
DB 118 VPRPGFGSPAPRKSKLICEARNFPTPKPTYSMLKQGLVBSGFTTDPVTIENKSTPOT 177
OY 167 ---KTLVSQLELDSDGTWTCTVLQONQKVEFKIDIVPCPAPEKSCDKHTTCPELLGG 222
DB 178 YKVISLTLSIEIDMLNLYVTCRV--DHRGLTFLKNVSTCAASPST-----DIL-- 225
```



```

FT TURN 226 229
FT TURN 231 232
FT STRAND 236 242
FT STRAND 252 252
FT STRAND 255 257
FT TURN 258 259
FT STRAND 260 260
FT TURN 265 266
FT STRAND 271 274
FT TURN 275 276
FT STRAND 277 285
FT HELIX 288 293
FT TURN 294 294
FT STRAND 297 302
FT TURN 304 305
FT STRAND 310 314
FT STRAND 322 322
FT STRAND 325 329
FT HELIX 334 337
FT STRAND 340 348
FT STRAND 350 351
FT STRAND 356 360
FT TURN 366 369
FT STRAND 374 374
FT STRAND 377 378
FT TURN 380 381
FT STRAND 384 385
FT STRAND 387 393
FT HELIX 394 398
FT TURN 399 400
FT STRAND 404 408
FT STRAND 418 421
SQ SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;

Query Match 13.4%; Score 361.5; DB 1; Length 428;
Best Local Similarity 28.9%; Pred. No. 4.5e-16;
Matches 102; Conservative 53; Mismatches 137; Indels 61; Gaps 10;

QY 90 GNPPLIIKLIKEDSDTYICEVEDQKEVQLVFGILTAN--SDTHLLQGQSLTLTLSPSP 147
DB 125 GHFPPITQL-----CLVSGYTPGTINITWLEDQVMDVDL----- 160
QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWTCTVLQNGKVEFKIDVPCAPRP 207
DB 161 ----STAISTQBELASTOSTGSELTLSQKHLSDRTTYCQVLYGHTPE----- 203
QY 208 KSCDKHTCTPEL-LGSPSVFLPPPKRDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDG 266
DB 204 ----DSTFKKADSNPREVSAVLSRSPFD-LFIKSPITICTLVVDLAPSKGTVALTWSPAS 259
QY 267 VEVHNAKTKPREQVNSTYRVSVLTVLHODWLNGEKYCKKVNKALPAPIETKISKAG 326
DB 260 GKVNHSTKEKEKQNGTLVTSTLPLVGTRDWEGETYQCRVTHPHLPALMKSTTKTSG 319
QY 327 QREPPQVYTL-----PPSRDELTKQNVSLTCLVKGYPYDIAVEMESN--GQENNYKTT 379
DB 320 PRAAPVYVAPNPWEGSRDKRT-----LAQLQNPEDPISVQWLNHEVQLPDAHSTT 374
QY 380 PVLVDSDGSFFLYSKLTVDKSRWQGNVSCSVMAHMLNHYT-QKSLSLSPG 431
DB 375 QPKTKGSGFFVSRLEVTAPAEWQKDEFCRAVHEDASQIVQAAVSINFG 427

RESULT 43
MUC_RABIT STANDARD; PRT; 458 AA.
AC P03988;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 1g mu chain C region secreted form.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RP [1]
RA SEQUENCE FROM N.A. (A2 ALLOTYPED).
RX MEDLINE=8408930; PubMed=6418803;
RA Bernstein K.E., Alexander C.B., Reddy E.P., Mage R.G.;
RT "Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain
of Vha2 allotype: comparisons with Vha1 and membrane mu sequences.";
RL J. Immunol. 132:490-495(1984).
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=During differentiation, B lymphocytes switch from
CC expression of isoform Membrane-bound to isoform Secreted;
CC Name=Secreted;
CC IsoId=P03988-1; Sequence=Displayed;
CC Name=Membrane-bound;
CC IsoId=P04221-1; Sequence=External;
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; K01357; -; NOT_ANNOTATED_CDS.
DR PIR; A02164; MHRB.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IGCL; 2.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
KW NON_TER
FT DOMAIN 1 106 CH1.
FT DOMAIN 107 222 CH2.
FT DOMAIN 223 327 CH3.
FT DOMAIN 328 458 CH4.
FT DISULFID 14 14 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFID 28 90 BY SIMILARITY.
FT DISULFID 137 200 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 219 219 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 249 308 BY SIMILARITY.
FT DISULFID 296 296 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 356 418 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 458 AA; 49897 MW; F338D6A3B28E4864 CRC64;

Query Match 13.3%; Score 360; DB 1; Length 458;
Best Local Similarity 23.8%; Pred. No. 6.2e-16;
Matches 118; Conservative 76; Mismatches 145; Indels 156; Gaps 22;

QY 21 AATGKNTVILGKKGVTELCTASQ--KKSIOFHW--KSNQIKILGNGSFLTKGPSKL 76
DB 17 ALTDGVLVAMG-----CLARDPLSSVTFMSFKNSRI----- 50
QY 77 NDRADSRSLMDQGNFLLIKLK-----IEDSDTY-ICEVEDQKEVQL 121
DB 51 -----SSRTV---RIFPVVKRGDKYMATSQVLVPSKDVQGTBEYLVCKVQHSNRRDLR 102
QY 122 VGLTANSTHLLQGQSLTLTLSPGSSPSVQCR-----SPR-- 159

```

```

Db 103 V-----SFEVDSELPNNVSVPFIPRDSFSGSCTRKSRLLCQATGFSPKQI 147
Qy 160 -----GKNIOGG-----KTLVSQLELDQSGT-----TCVY-----L 187
Db 148 SVSWLRDGGKVSGLVTKPEAEYKAGAPATPSISSMLTTESDWLSSGLTYCRVDHGI 207
Qy 188 QNOKVEFKIDIVPCPAPEPKSCDKHTPCRELGGSPVFLPPPKDKDTLMSRTPEVTCV 247
Db 208 FPDKNVSSSECSSTPSF-----GIQVFPILPSPADT-FLSKSARLLCL 250
Qy 248 VVDVSHEDPEVKFNNYVDGVEVNAKT-----KPREQYNSTYRVSVLTVLHODMLNGK 302
Db 251 VTDLTYYG-SLNISSW-----ASHNGKALDTHMNTESHNPATFSAMGASVCAEDMESGE 304
Qy 303 EYKCVSNKALPAPLEKTSKAKQPRE-POVYTLPPSRDEL-KNOVSLTCLYKGFPS 360
Db 305 QCTCTVTHADLPFLKHTISKREYAKHPAVYVLPABEQDLRESATVCLVGFSPDA 364
Qy 361 DIAVESMGQP--ENNYKTPPVLD--SDGSFPLYSKLTVDKSRMOQGNVSCSVMEHA 416
Db 365 DVFGVMQGRGQPLSSDKVTSAPAPEPQAPGLYFTHSTLTYEBDMNGSETTTCVVGHEA 424
Qy 417 LHNHYTQKSLSLSPG 431
Db 425 LPHWVTERTVDKSTG 439

RESULT 44
MUC_SUNMU
ID MUC_SUNMU STANDARD; PRT; 457 AA.
AC P20768;
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig mu chain C region.
OS Suncus murinus (House shrew) (Musk shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Soricidae; Crocidurinae; Suncus.
ON NCBI_TaxId=9378;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89232144; PubMed=2497033;
RA Ishiguro H., Ichihara Y., Namikawa T., Nagatsu T., Kurosawa Y.;
RT "Nucleotide sequence of Suncus murinus immunoglobulin mu gene and
RL comparison with mouse and human mu genes.";
PUBS Lett. 247:317-322(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X13920; CAA32113.1; ALT_INIT.
DR PIR: S03961; S03961.
DR HSR: F01842; 7FAB.
DR InterPro: IPR007110; IG-1like.
DR InterPro: IPR003597; IG_C1.
DR InterPro: IPR003006; IG_MHC.
DR Pfam: PF00047; Ig_4.
DR SMART: SM00407; IGc1_3.
DR PROSITE: PS00835; IG_LIKE; 4.
DR PROSITE: PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON TER 1 105 CH1.
FT DOMAIN 106 220 CH2.
FT DOMAIN 221 326 CH3.
FT DOMAIN 327 457 CH4.

```

```

FT DISULFD 13 13 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFD 27 89 BY SIMILARITY.
FT DISULFD 136 200 BY SIMILARITY.
FT DISULFD 217 217 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFD 248 307 BY SIMILARITY.
FT DISULFD 295 295 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFD 355 417 BY SIMILARITY.
FT CARBOHYD 456 456 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 457 AA; 50074 MW; 56C8C086D5A4462E9 CMC64;

Query Match 13.1%; Score 354.5; DB 1; Length 457;
Best Local Similarity 25.5%; Pred. No. 1,46-15;
Matches 122; Conservative 71; Mismatches 163; Indels 123; Gaps 19;

Qy 19 LPAATQGNKVVLGKKGDTVELTCTASQ--KKSIOFHMKNQIKILNQSGSFLTKGPEKL 76
Db 17 LPDETQ-----VTLGCLARDFLPRTVTFMKKNSSSI----- 49
Qy 77 NDRADSRSLWDQGNFPLIKNLK-----IEDSDTYI--CEVEDQKEVOL 121
Db 50 -----SQGNIT--NPEDEVFTGKVMATSGVLPSTALIGSTDYITCTHTGTTGEKK 101
Qy 122 VFGILANSDTHTLLOGSILTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDQSG- 180
Db 102 V-----ELQVTPPELPNNVSIFV--PPR--NSFGNHPRTSQLCOASGF 141
Qy 181 ---TWCTVLONOKKVEFKIDIVPCPAPEPKSCDT-----HICP 217
Db 142 SPRTIWSWLRGEPVQPSLSTSAVEAEFGSGPTTFRVYSLRTITENEWLSQREFTCQ 201
Qy 218 EL-----LG-----GPSVFLPPPKDKDTLMSRTPEVTCVVVDVSHEDPEV 258
Db 202 ALHKGILFGQKNVSVCMGDDTSTGISVFLPPTFAN-IFLQNSQTLCLVGLATYD-SL 259
Qy 259 KFNMYVDGVEVNAKTPREEQYNSTYRVSVLTVLHODMLNGEKYKCKSNKALPAPLE 318
Db 260 DISMRONGEALQTHVNISESHPNSTFTAKGASVCEBESGESGFTCTVGHSDLPFLK 319
Qy 319 KTIKAKGQPRE-POVYTLPPSRDEL-TKNOVSLTCLYKGFPSDIAVESMGQP--EN 374
Db 320 QSLSRPDVANDPSPSVLPAPQEQCLKRESASITCLVKDPSPPDVFGVOMHHGQPDPEK 379
Qy 375 NYKTPPVLD--SDGSFPLYSKLTVDKSRMOQGNVSCSVMEHAALHNHYTQKSLSLSPG 431
Db 380 HYVTSNPTPEQNGFLVFHSILTVSEKDMSSGSGFCVGHALLPLSTVEKAVDKTSG 438

RESULT 45
MUC_MESAU
ID MUC_MESAU STANDARD; PRT; 454 AA.
AC P06337;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE Ig mu chain C region.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
ON NCBI_TaxId=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=65297761; PubMed=2994005;
RA McGuire K.L., Duncan W.R., Tucker P.W.;
RT "Phylogenetic conservation of immunoglobulin heavy chains: direct
RL Nucleic Acids Res. 13:5611-5628(1985).

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on items
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).

CC -----

DR EMBL; X02804; CAA26574.1; -
DR HSSP; P01854; 1IGE.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.

FT NON_TER 1 1
FT DOMAIN 1 105 CH1.
FT DOMAIN 106 218 CH2.
FT DOMAIN 329 324 CH3.
FT DOMAIN 325 454 CH4.
FT DISULFID 13 13 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFID 27 88 BY SIMILARITY.
FT DISULFID 135 198 BY SIMILARITY.
FT DISULFID 215 215 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 245 304 BY SIMILARITY.
FT DISULFID 292 292 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 352 414 BY SIMILARITY.
FT DISULFID 453 453 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT CARBOHYD 45 45 N-LINKED (GLCNAc . . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAc . . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAc . . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAc . . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAc . . .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAc . . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAc . . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAc . . .) (POTENTIAL).
SQ SEQUENCE 454 AA; 50196 MW; 4DA2134612BE1469 CRC64;

Query Match 12.7%; Score 344; DB 1; Length 454;
Best Local Similarity 24.9%; Pred. No. 6; 4e-15;
Matches 115; Conservative 74; Mismatches 158; Indels 114; Gaps 19;

QY 37 VELTTSAG--KKSGQFMFKKSNQKILGNQ--SFLTKGFSKLNDRADSRSLMDQGNF 92
DB 23 VAMGGLADFLPSSISFSMNYONKSEV--NOGVTFPL-----RMGEKATATSVQVFL 73
QY 93 PLIKNLKLEIDSDPY-ICEVEDQKEVQLVFGLSANDTHLQGSLLTLESPPGSSP 151
DB 74 P---PKSVLEGSDEYLCKVHNGNTKDLRV-----PIPVTEMNPVNSV 115
QY 152 SVQCR-----SPR-----GKNIQG----- 166
DB 116 FVPSRDASGPAKRSRLFCEASNSNPQGITVSMYLRDQKPYKSGTTEPVTPEDRGSGPR 175
QY 167 -----KTLVSQLELDSDGWTCTVLYNQKVEFKIDIVPCRAPEKSCDHTTCEBLLG 221
DB 176 TYKVLSTLTITTESDWLNLNVYTCRV--DHRGLTFMKNVSSSTCAASPST----- 221
QY 222 GPSVLEFPKPKEDT-LMTSRPELVNVVDVSHEDPEVKFMVYNDGVEVNHAKTPEBEQ 280
DB 222 --DIDAFPIPSFGVGIPLNKSATLTCLVNTNLATYD-TLNTSSSGSGEPLETKTLCLES 278
QY 281 YNSTYRVSVLTVLHQMVLNGKEVCKSNKALPARIETKISKAGQERE-----PQYUT 335
DB 279 PNGTSAIGAEANVCVEDMDSGKEPFCVYTHHDLRPEQCKFKISK-----PRENNKTPRAYQ 334
QY 336 LPSPADEL-TKNQVSLTCLVKGFGYSDIAVEMESGQR-ENNYKTPRPVLDSDGS--FF 390
DB 335 QPLAREQILLESASVTVCLVKGFGSPADIFVQMLGQGLSDQKYTASPKMEBQAPAHLYF 394

Gy		391	LVSKLTVDSRMQGGVSCSYWHEALHNHYRQKSLSPG	431
Dd		395	THSVLVTVEEWNMSGETTCVGHEALPMYTERTVDRSTGT	435
RESULT 46				
ID	_CD7_HUMAN	STANDARD;	PRT;	240 AA.
AC	P09564;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	10-MAR-1989 (Rel. 10, Last sequence update)			
DE	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	T-cell antigen CD7 precursor (GP40) (T-cell leukemia antigen) (TP41) (Leu-9).			
GN	Cd7.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	[1]			
RX	SEQUENCE FROM N.A., PubMed=3501369;			
RX	Aurifio A., Seed B. ;			
RT	"Molecular cloning of two CD7 (T-cell leukemia antigen) cDNAs by a			
RT	COS cell expression system."			
RL	EMBO J. 6:3313-3316(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A., PubMed=1703303;			
RX	Medline=91110576;			
RA	Schanberg L.E., Fluenor D.B., Kuritzberg J., Haynes B.F., Kaufman R.E.,			
RT	"Isolation and characterization of the genomic human CD7 gene:			
RT	structural similarity with the murine Thy-1 gene."			
RT	Proc. Natl. Acad. Sci. U.S.A. 88:603-607(1991).			
RN	[3]			
RP	SEQUENCE FROM N.A., TISSUE=Muscle;			
RX	Medline=22388257; PubMed=12477932;			
RA	Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hasteh F., Diachenko L., Marasina K., Farmer A.A., Rubin G.W., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Udell T.B., Toehlyuki S., Carninci P., Prange C.C., Rana S.S., McQuillan N.A., Peters G.J., Abramson R.D., Mullihay S.J., Bosak S.A., Coleman P.J., McKernan K.J., Malek S.A., Gunaratne P.H., Richards S., Wolcy K.C., Hale S., Garcia A.M., Gay L.J., Huljck S.W., Villalón D.K., Munzy D.M., Sodegren E.J., Lu X., Gibbs R.A., Fahney J., Holton B., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield J.S.N., Krzywinski M.I., Skalska U., Smialowski D.E., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002). [4] SEQUENCE OF 205-240 FROM N.A. MEDLINE=91267564; PubMed=17111009; Yoshikawa K., Seto W., Ueda R., Obata Y., Notake K., Yokochi T., Takahashi T.; "Molecular cloning of the gene coding for the human T cell differentiation antigen CD7." Immunogenetics 33:352-360(1991). [5] TOPOLOGY. MEDLINE=90063052; PubMed=2479685; Ware R.E., Seearce R.M., Dietz M.A., Stamer C.F., Palke T.J., Haynes B.F.; "Characterization of the surface topography and putative tertiary structure of the human CD7 molecule."			

Matches 113; Conservative 67; Mismatches 155; Indels 110; Gaps 17;

QY 23 TQGNKY---VLGKGGDTVELTCAQKKSIGPHMKNSNOIKILNGSGFLTGPSKLNDR 79
 DB 49 TGLKTYPSVLNKKGT-----YTRSSQULTTES----VGSSTKYCE 86
 QY 80 ADSRSLMDQGNFPLIIRKLIKEDSDTYICEVEDQKEVQLVFGLTANSDTL----- 133
 DB 87 VARGESLW-----IKELIDCKGDIYPPVILITSSSEITRRRA 126
 QY 134 -----LOGSLTLT-----LESPPSSPSVOCRSRPRGNIGKGTLSVSOQLD 178
 DB 127 TVLCSTIDPHPSBITSMKDGQPMDSGFVSTPCVNG---NFSATSLTLVPAGEWPS 182
 QY 179 SGTMTCTVLQONQKVFEXI--DIVPCRPAPKSCDHTHCPELLGSPSFLPPKPKDYL 236
 DB 183 NTVYTQVHAQETQSRNITGSOVPS-----IGDPVTKLPPSTIEQVL 226
 QY 237 MISRPETVCVVVDVSHEDPEVFNMYVDGVEVHNAKTRPREQY-----NSTYRVSVL 291
 DB 227 -LEATVTLTCV---VSNAPYGVNVSM-----TQEKPLKSEIAYQGEBSDSVISTY 274
 QY 292 TVLHODWLNGKEYKCVSNKALPAPIKTIKSKAKGP-REPOY-YTLPPSRDELTKQVS 349
 DB 275 NISTQAMLSGAEFYCVVSHQDLPTPLRASIHKKEVNDLREPFVSVLLPPAEDVSAQRFLS 334
 QY 350 LTVLVGFPSDIAVEMESNGOPEN--NYKTPPVLDSCG-SFELYSKLTVDKSRQOQN 406
 DB 335 LTVLVGFSFRELFIKMTVNDKSVNGNYKTEVMAENDNRSPFIYSLLSIAAEWASGA 394
 QY 407 VFSCSVMEALHNHYTKSLSLSPG 431
 DB 395 SYSCVGVHEAIPLKINRTVKNSSG 419

RESULT 48
 MUC_CHICK
 ID MUC_CHICK STANDARD; PRT; 446 AA.
 AC P01675;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig mu chain C region.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauri; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN (1)
 RP SEQUENCE OF 80-446 FROM N.A.
 RX MEDLINE=63299221; PubMed=6310496;
 RA Dahan A., Reynaud C.-A., Weill J.-C.;
 RT "Nucleotide sequence of the constant region of a chicken mu heavy
 chain immunoglobulin mRNA."
 RL Nucleic Acids Res. 11:5381-5389(1983).
 RN (2)
 RP SEQUENCE OF 1-79 FROM N.A.
 RA Weill J.-C.;
 RL Submitted (MAR-1986) to the EMBL/Genbank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X01613; CAA25762.1; --
 DR PIR: A02170; MHCH.
 DR HSSP: P01857; IFCL.
 DR InterPro: IPR007110; Ig_1like.
 DR InterPro: IPR003597; Ig_c1.

DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig_3.
 DR SMART: SM00407; IgC1. 1.
 DR PROSITE: PS00835; IG_LIKE. 4.
 DR PROSITE: PS00290; IG_MHC; 3.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON TER 1 105
 FT DOMAIN 1 105 CH1.
 FT DOMAIN 106 209 CH2.
 FT DOMAIN 210 316 CH3.
 FT DOMAIN 317 427 CH4.
 FT DOMAIN 428 446 C-TERMINAL REGION.
 FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN)
 FT DISULFID 27 85 (BY SIMILARITY).
 FT DISULFID 135 190 BY SIMILARITY.
 FT DISULFID 237 296 BY SIMILARITY.
 FT DISULFID 284 284 INTERCHAIN (WITH A HEAVY CHAIN)
 FT DISULFID 344 406 (BY SIMILARITY).
 FT DISULFID 445 445 INTERCHAIN (WITH A HEAVY CHAIN)
 FT CARBOHYD 45 45 (BY SIMILARITY).
 FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 158 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 446 AA; 48173 MW; 3CB0CE108949BD17 CRC64;

Query Match 11.5%; Score 310.5; DB 1; Length 446;
 Best Local Similarity 26.2%; Pred. No. 8.7e-13;
 Matches 119; Conservative 67; Mismatches 179; Indels 89; Gaps 19;

QY 28 VLGKGGDTV-ELTCTA--SOKKSIQPHMKNSNOIKILNGSGFLTGPSKLNDRASRR 84
 DB 13 LVLCSPSBSVTVYGCAPDFQPSSTIAFTWPSNNSSVSGM--DIVPKYISGPPYAVSR- 69
 QY 85 SLMDQGNFPLIIRKLIKEDSDTYICEVEDQKEVQLVFG---LTANSDTLQOQSLT 140
 DB 70 -----IQNQGSEKKEKQPFRCRAHPRGNVSVNPPPIPTPN-----GIPLF 113
 QY 141 LTLESP-----PSSSPVQC-----RSPR-----GKXIQQCK 167
 DB 114 VTMHPSPSEDEFGPPRNASILCQTRGRRRPTEVTYKNGSPVAAAATTATTVGPPVARS 173
 QY 168 TLSVSOLELDQSGTCTVYLQNKQKVEFKIDIVPCPAPEPKSCDHTHC---PELLGPS 224
 DB 174 RISVTESEWDGATFSCV-----EGEMRNTSKRMCGLEPPVQODIA 216
 QY 225 VLFPPPKDRLTMSRTEVTCVVVDVSHEDP-EVKFNMYVDGVEVHNAKTRPREQVNS 283
 DB 217 IRVIRPSPVD-IFISKATLTCTRVSNMVAADGLEWS- WKEKGGLTETALGK-RVLQNSG 273
 QY 284 TYRVSVTLVTHQDLNGKEYKCVSNKALPAPIKTIKSKA-GOPREPOVYTLPPSRDE 342
 DB 274 LYTVDGVAIVCASEWDGDDGYCVKVNHPDLLFPMEBKRRKTKASNARPPSVVFPPTQ 333
 QY 343 LTKNO-VSLTCLVKGFPYSDIAVEMESNGO--PENNYTTTPVLDSD--DGSFELYSKLT 397
 DB 334 LKNGRLSVTCQAQGFNPPLFLFRMRNGEPLPQOSVTSAPMAENPENESTVAIVSLGV 393
 QY 398 DKSRMQQGNVFSQSVMEALHNHYTKSLSLSPG 431
 DB 394 GAEEWGAGNVYTCVLGHALPLQLAQKSVDRASG 427

RESULT 49
 HCM_HETFR
 ID HCM_HETFR STANDARD; PRT; 461 AA.
 AC P23088;
 DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig heavy chain C region, membrane-bound form (Clone 3050).
 OS Heterodontus francisci (Horn shark).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphi; Heterodontidae; Heterodontiformes;
 OC Heterodontidae; Heterodontus.
 NCBI_TaxId=7792;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=88328985; PubMed=3138109;
 RA Kokubu F., Hinds K., Litman R., Shambloot M.J., Litman G.W.;
 RT "Complete structure and organization of immunoglobulin heavy chain
 RT constant region genes in a phylogenetically primitive vertebrate.";
 RL EMBL J. 7:1979-1988(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, X07781; CAA30614.1; ALT_SEQ.
 DR PIR, S01854; HVRKCO.
 DR HSRP, P01842; 7BAB.
 DR InterPro, IPR007110; Ig_1-like.
 DR InterPro, IPR003597; Ig_c1.
 DR InterPro, IPR003006; Ig_MHC.
 DR Pfam, PF00047; Ig_4.
 DR SMART, SM00407; IGc1_2.
 DR PROSITE, PS00835; IG_Like; 4.
 DR PROSITE, PS00290; IG_MHC; 3.
 DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW Alternative splicing; Transmembrane.
 FT NON_TER 1 1
 FT DOMAIN 1 99 CH1.
 FT DOMAIN 100 205 CH2.
 FT DOMAIN 206 308 CH3.
 FT DOMAIN 309 418 CH4.
 FT TRANSMEM 438 458 POTENTIAL.
 FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 415 415 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 461 AA; 50762 MW; 029E91C8B5D5911 CRC64;
 Query Match 11.2%; Score 301.5; DB 1; Length 461;
 Best Local Similarity 26.7%; Pred. No. 3.4e-12;
 Matches 121; Conservative 67; Mismatches 161; Indels 105; Gaps 23;
 QY 53 WKNSNQIILNGSGFLTKGPKNDKDRSRSLMDQGNFLITKLNKIEISDYICV- 111
 DB 41 WKQDNEPITTG---LNTYPSVLNKKKGTYS---SQTITSESVSSSKTY-CEVR 88
 QY 112 -----EDQKEEVQLLVFG--LTANSDFTLGOSITLTL- 142
 DB 89 RGESEVWKEIDPCCKDKXHPVITVLQSSSELTSRKATVLCSTIDFH---PESTIVSWL 145
 QY 143 -----LESPGSSPSVQGRSPRKN-1QGKTLISVSOLELQDSGTWTCTVLONOKYEF 195
 DB 146 KDQGMESGFTSPT--C-----GVNGTSATSRLTVPARFWTKNVYTCQV--SHQGVQ 197
 QY 196 KIDI-----VPCPAPRPSKCDKTHTCPELLGGPSVTLFPKPKDITLMISRETVTCVVDV 251
 DB 198 SRNITGSGVPC---SCN-----DPVILKLPSPISIEQVL-LEATVTLTCTV---V 237

QY 252 SHEDPEVKFNYYDGVENHNAKTKPREQYNSYRVVSVLTVLHQDMLNGEKYCKVSNK 311
 DB 238 SNAHYGVNVS---TQOKSLKSEIAVQPGEDADSVISTVNIISQANLSCGFPCVANHQ 294
 QY 312 ALPAPIKTIISKAKQP-REPVQ-YTLPSRDELTKQVSLTCLVKGFPSPDIWEMESN 369
 DB 295 DLPLPLASIHKEVKDLREPSVSIILSPADVSAQFLSLTCLVRFPPREIVKMTVN 354
 QY 370 QQPEN--NYKTPPVLDSD--GSFLVSKLTVYDKSRKQGVFSGSWHEALHNHYTKSL 426
 DB 355 DKSVNPGNVKNTVEWANDNSYFIYSLSTIAEEMASGASVSCVGHETL----- 405
 QY 427 SLSPGLQDETCAEQDGLDGLTTPDPAPASL 460
 DB 406 ---PLKINRTVNTKSSDS--DHWIEDNEESAI 435
 RESULT 50
 AL22 HUMAN STANDARD; PRT; 340 AA.
 ID AL22 HUMAN
 AC P01877;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ig alpha-2 chain C region.
 GN IGH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84130179; PubMed=6421489;
 RA Flanagan J.G., Lefranc M.-P., Rabbitts T.H.;
 RT "Mechanisms of divergence and convergence of the human immunoglobulin
 RT alpha 1 and alpha 2 constant region gene sequences.";
 RL Cell 36:681-688(1984).
 RN [2]
 RP SEQUENCE (BUT).
 RX MEDLINE=78137069; PubMed=416441;
 RA Torano A., Putnam F.W.;
 RT "Complete amino acid sequence of the alpha 2 heavy chain of a human
 RT IgA2 immunoglobulin of the A2m (2) allotype.";
 RL Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).
 RN [3]
 RP SEQUENCE (MYELOMA PROTEIN LAN).
 RX MEDLINE=79180140; PubMed=286295;
 RA Teuszida Y., Wang C.-C., Putnam F.W.;
 RT "Structure of the A2m(1) allotype of human IgA -- a recombinant
 RT molecule.";
 RL Proc. Natl. Acad. Sci. U.S.A. 76:1104-1108(1979).
 RN [4]
 RP REVIEW.
 RX MEDLINE=91054387; PubMed=2241915;
 RA Kerr M.A.;
 RT "The structure and function of human IgA.";
 RL Biochem. J. 271:285-296(1990).
 CC -1- FUNCTION: Ig alpha is the major immunoglobulin class in body
 CC secretions. It may serve both to defend against local infection
 CC and to prevent access of foreign antigens to the general
 CC immunologic system.
 CC -1- SUBUNIT: Monomeric or polymeric.
 CC -1- MISCELLANEOUS: The sequence of the A2m(1) allotype is shown.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----


```
DR EMBL: J00221; AAB59396.1; ALT_INIT.
DR PIR: B2360; B2360.
DR HSSP: P01810; 2FBJ.
DR GeneW: HGNC:5479; IGHA2.
DR GO: GO:0005624; C:membrane fraction; NAS.
DR GO: GO:0003823; F:antigen binding; TAS.
DR GO: GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig-cl.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; IgC1; 2.
DR PROSITE: PS00835; IG_LIKE; 3.
DR PROSITE: PS00290; IG_MHC; 2.
KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
FT NON_TER 1 98
FT DOMAIN 112 207
FT DOMAIN 215 317
FT DISULFID 26 85
FT DISULFID 101 101
FT DISULFID 109 109
FT DISULFID 110 167
FT DISULFID 134 191
FT DISULFID 169 169
FT DISULFID 179 179
FT DISULFID 237 300
FT DISULFID 339 339
FT CARBOHYD 47 47
FT CARBOHYD 92 92
FT CARBOHYD 131 131
FT CARBOHYD 205 205
FT CARBOHYD 327 327
FT VARIANT 93 93
FT VARIANT 102 102
FT VARIANT 279 279
FT VARIANT 296 296
FT VARIANT 326 326
FT VARIANT 335 335
SQ SEQUENCE 340 AA; 36508 MW; 98922700756F3276 CRC64;

Query Match 10.9%; Score 294.5; DB 1; Length 340;
Best Local Similarity 29.2%; Pred. No. 6.5e-12;
Matches 103; Conservative 50; Mismatches 143; Indels 57; Gaps 18;

QY 122 VFGLTANSDTH-----LQG-----QSLLTLESPPGSSPVSQCRS-PRGNIGGKTK 168
DB 8 VFPLSLDSTFGDGNVYVACLVGPFQPEPLSTVWSE---SGQVNTARNPPSPDAGDLY 64
QY 169 LSVSGL-----QDSGTWTCTVLQNKQKVEFKIDI-VPCPAPEPKSCDKTHTCPBELGG 222
DB 65 TTSSQTLTPATGCPDGKSVTCHVKHYTNPSQ---DVTVCPPVPPPPC-----C-----H 111
QY 223 PSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVNAKTPREEQVN 282
DB 112 PRSLRHPALPD--LLGSEANLICTTLGL--RDASGATFTWTPSSGK--SAVQGPPEPRLDC 167
QY 283 STYRVVSVLTVLTHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDE 342
DB 168 GCSVSVSIVLPSCAQPNHGETFTCTAAHPELKTPLTANITKS-GNTRPREVHLPPPSSE 226
QY 343 LTKNQ-VSLTCLVKGFPYSDIAEWESNQ--PENNYKTPPVLD--SDG--SFFLYSKLT 396
DB 227 LALNELVTLTCLARGFSKDVLRVWLQSGQELPREKYLTVASRQSPSGTGTTPAVTSLR 286
```

```
QY 397 VDKSRQGNVFSQSYMEALHNHYTQKSLSPG-----LQDFTC 438
DB 287 VAAPDKKDDTSCVNGHEALPLATFQTKTIDLAKGPTHVNVSVYAAVDGTC 339

RESULT 51
HVC5_HETFR
ID HVC5_HETFR STANDARD; PRT; 438 AA.
AC P23087;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain C region, secreted form (Clone 3050).
OS Heterodontus francisci (Horn shark).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Heterodontidae; Heterodontiformes;
OC Heterodontidae; Heterodontus.
OX NCBI_TaxId=7792;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=脾脏;
RX MEDLINE=85328985; PubMed=3138109;
RA Kokubu F., Hinds K., Litman R., Shambloot M.J., Litman G.W.;
RT "Complete structure and organization of immunoglobulin heavy chain
constant region genes in a phylogenetically primitive vertebrate.";
RL EMBO J. 7:1979-1988(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: X07781; CAA30613.1; ALT_INIT.
CC DR PIR: S01853; HVRKCS.
CC DR HSSP: P01842; 7FAB.
CC DR InterPro: IPR007110; Ig-like.
CC DR InterPro: IPR003597; Ig-cl.
CC DR InterPro: IPR003006; Ig_MHC.
CC DR Pfam: PF00047; Ig; 4.
CC DR SMART: SM00407; IgC1; 2.
CC DR PROSITE: PS00835; IG_LIKE; 4.
CC DR PROSITE: PS00290; IG_MHC; 3.
KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1 99
FT DOMAIN 100 205
FT DOMAIN 206 308
FT DOMAIN 309 418
FT CARBOHYD 164 164
FT CARBOHYD 200 200
FT CARBOHYD 245 245
FT CARBOHYD 275 275
FT CARBOHYD 374 374
FT CARBOHYD 411 411
FT CARBOHYD 415 415
FT CARBOHYD 425 425
SQ SEQUENCE 438 AA; 47920 MW; 9BDF37C3BBD7D15B CRC64;

Query Match 10.9%; Score 294; DB 1; Length 438;
Best Local Similarity 27.1%; Pred. No. 9.6e-12;
Matches 115; Conservative 65; Mismatches 153; Indels 92; Gaps 21;

QY 53 WKNSNDIKILNGQSLTGPESKLDNRASRLMDQGNFPLINLKIKEDSDTYICEV- 111
DB 41 WKDNEPITTG-----LKTYPVILNKGTYTDS-----SOLLTLESSESSKIY-CEVR 88
QY 112 -----EDQKEEVQLVFG--LTVNSDTHLQGSLLTLT-- 142
DB 89 RGESWIKELIPDCKDKVHPVILVLTQSSEELTSRFAVLCSIIDFH---PESTIVSWL 145
```

[illegible]

RESULT 52	ALC_RABIT	STANDARD:	PRT:	299 AA.
AC	ALC_RABIT			
AC	P01879;			
DT	21-JUL-1986	(Rel. 01, Created)		
DT	21-JUL-1986	(Rel. 01, Last sequence update)		
DT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Ig alpha chain C region (Fragment).			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84144059; PubMed=6322114;			
RA	Knight K.L., Martens C.L., Stoklosa C.M., Schneiderman R.D.;			
RT	"Genes encoding alpha-heavy chains of rabbit IgA: characterization of			
RT	cDNA encoding IgA-g subclasses alpha-chains."			
RL	Nucleic Acids Res. 12:1657-1670(1984).			
CC	-1- FUNCTION: Ig alpha is the major immunoglobulin class in body			
CC	secretions. It may serve both to defend against local infection			
CC	and to prevent access of foreign antigens to the general			
CC	immunologic system.			
CC	-1- MISCELLANEOUS: This immunoglobulin belongs to the IGA-G subclass.			
CC	It was isolated from a rabbit homozygous FOR A2, N80, DE12, 15,			
CC	F71, G75 heavy chain haplotype.			
CC	-1- SIMILARITY: Contains 2 immunoglobulin-like domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X00353; CA825100.1; -			
DR	PIR; A02174; AHRB.			
DR	HSSP; P01879; 1FC1.			
DR	InterPro; IPR007110; Ig-1Ike.			
DR	InterPro; IPR003597; Ig_C1.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	Pfam; PF00047; Ig_2.			
DR	SMART; SM00407; IGC1_2			
DR	PROSITE; PS50835; IG-LIKE; 2.			
DR	PROSITE; PS00290; IG_MHC; 2.			
FW	Immunoglobulin domain; Immunoglobulin C region; Repeat.			

FT	NON TER	1	1						
FT	DOMAIN	71	167	IG-LIKE 1.					
FT	DOMAIN	174	276	IG-LIKE 2.					
SQ	SEQUENCE	299 AA;	32256 MW;	2512FE3F62E9A223 CRC64;					
Query Match		10.8%:	Score 293;	DB 1;	Length 299;				
Best Local Similarity		31.7%:	Pred. NO. 6.9e-12;						
Matches 78; Conservative		42;	Mismatches 106;	Indels	20;	Gaps	10;		

Qy 200 VPCPAEPKSCKHTCP-----ELLGG--SVSLPPPKMDTMIIRTEVTCVVVDV 252
Db 41 LPVPFDFCCPANSCTCPSSSSRNLIISCGQSLSLQRPDLG-LLLRDMSLTCTLSGLK 99
Qy 253 HEDPEVKFNWYDGVENVNAKTPREEOYNSTYRVSVLTVLQDWLNGEKYCKYSNKA 312
Db 100 NPEDAV-FTW--EPNCGNBPVQQRQORDLSGCVSSVSLPSAETWAKAREFTCTVTHPE 156
Qy 313 LPA-PIEKTIISKAKQPREPOVYTLPRERDLTGN-QVSLTCAVKGFPYPRDIANVESNG 370
Db 157 IDSGSLTATISIGVTP--POVHLTPRESEELALBOVTLTCLVGRGSPKDVLTSMNHQ 214
Qy 371 Q--PENNY---KTPRPVLDSGSPFLYSKLTIVDKSRWQOQNVFSCSYMHALNNHTOKS 425
Db 215 QEVPEDSFLWVMSKBPSSQDKATAYATLTSLAVPAEDMNQDPTYSQVHGHEGLAHEFTQK 274
Qy 426 LSLSPG 431
Db 275 IDRLAG 280

RESULT	53
AC	P23084;
DT	01-NOV-1991 (Rel. 20, Created)
DT	01-NOV-1991 (Rel. 20, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Ig heavy chain C region (Clone 6125) (Fragment).
OS	Heterodontus francisci (Horn shark).
OC	Euxaryotes; Metazoa; Chordata; Craniota; Vertebrata; Chondrichthyes;
OC	Elasmobranchii; Galeomorphii; Heterodontoidae; Heterodontiformes;
OC	Heterodontidae; Heterodontus.
ON	NCHI_Taxid=7792;
OR	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=68328985; Pubmed=3138109;
RA	Kokubu F., Hinds K., Litman R., Shambloct M.J., Litman G.W.;
RT	"Complete structure and organization of immunoglobulin heavy chain
RL	constant region genes in a phylogenetically primitive vertebrate.";
RL	EMBO J. 7:1979-1988(1988).
CC	- I SIMILARITY: Contains 3 immunoglobulin-like domains.
CC	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	
DR	EMBL; X07783; CAA30616.1; -.
DR	PIR; S01851; HVKCKS.
DR	HSSP; P01857; 1FCL.
DR	InterPro; IPR007110; Ig_1like.
DR	InterPro; IPR003597; Ig_c1.
DR	InterPro; IPR003006; Ig_MHC.
DR	Pfam; PF00047; Ig_1.
DR	SMART; SMO0407; Igcl; 2.
DR	PROSITE; PSS0835; IG_LIKE; 3.
DR	PROSITE; PSS00290; IG_MHC; 3.
DR	Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
TM	
TM	NON TER
TM	DOMAIN
TM	40 134
TM	1 1
TM	IG-LIKE 1.

```

FT DOMAIN 145 237 IG-LIKE 2.
FT DOMAIN 247 347 IG-LIKE 3.
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 370 AA; 40586 MW; 32833ABEFADDD0 CRC64;

Query Match 10.6%; Score 286.5; DB 1; Length 370;
Best Local Similarity 26.2%; Pred. No. 2.3e-11;
Matches 101; Conservative 59; Mismatches 155; Indels 71; Gaps 16;

QY 72 GRSKNDRADRSRLM-----DQGNFPLIKNLKIEDSD-----TYICEVEQKE 117
DB 11 GSKKICEVRGESLWIKELPDCKGIVHPTVILTQTSSEELTSRFAFVLCIIDFHE 70
QY 118 VOLVLEGLTAN--SDPHLLQGSLLTLSPSPSSPVOCSPRGKNIQSGKTLVSQLE 175
DB 71 -----AIVNMKDKDHPESEFVT-----SACETNG---NFSATSLTYPAE 111
QY 176 LODSGTWCTVLQNOKKVEFKIDIVCPAPEPKSCDKHTCPPELLGSPVFLPPPKDT 235
DB 112 WFTNTVYTCQVSHQATQSRNI---TGSPPDSECNH-----PAIKLLPPIEGV 157
QY 236 LMSRPEPTCVVDVSHEDPEYKFNWYVDGVEVNAKTKPREQY-----NSTYVAV 290
DB 158 L-LEAVTLTLCV---VSNAIPYGVNSW-----TQEQKPLKSEIAVOPGEDPDSVIST 205
QY 291 LTVLHODMNGKEYKCKVSNKALPAPIEKTISK-AGQPREPOVYTLPPSRDELTKNQ-V 348
DB 206 VDISQAMLSAEAVFCVYVSHQDLPTLRDSIHKEAKDLREPEVSULLPAEISMERL 265
QY 349 SLTCLVKGFFYPSDIAVESNGQPEN--NYKTPPYLDS-DGFFLYSKLTVDKSRWQ 405
DB 266 SLTCLVGRGSPREIFVKMTVNDKSNVPGNYKTEVMAENDKSFYISLISIAEEMWAG 325
QY 406 NVFSCSVMEALHNHTOKSLSPG 431
DB 326 ASYSCVGHGHAIPLKIINRTVKNSSG 351

RESULT 54
HVC3_HETFR
ID HVC3_HETFR STANDARD; PRT; 393 AA.
AC P23086;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
OS Ig heavy chain C region (Clone 6121) (Fragment).
OC Heterodontus francisci (Horn shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Heterodontidae; Heterodontiformes;
OC Heterodontidae; Heterodontus.
OC NCBI_TaxId=7792;
OX RN
RP SEQUENCE FROM N.A.
RX TISSUE=Spleen;
RX MEDLINE=88328985; PubMed=3138109;
RA Kokubu F., Hinde K., Litman R., Shambloot M.J., Litman G.W.;
RT "Complete structure and organization of immunoglobulin heavy chain
concentric region genes in a phylogenetically primitive vertebrate.";
RL EMBL J. 7:1979-1988 (1988).
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL: X07782; CAA30615.1; -
DR PIR: S01852; HVRK1.
DR HSRP: P01842; 7FRB.
DR InterPro: IPR007110; Ig-1-like.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00407; IgC1. 2.
DR PROSITE: PS00835; Ig_LIKE; 3.
DR PROSITE: PS00290; Ig_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
FT NON TER 1 1
FT DOMAIN 63 157 IG-LIKE 1.
FT DOMAIN 158 250 IG-LIKE 2.
FT DOMAIN 270 370 IG-LIKE 3.
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 393 AA; 43081 MW; 4E44D076972F18B5 CRC64;

Query Match 10.4%; Score 281; DB 1; Length 393;
Best Local Similarity 26.4%; Pred. No. 5.7e-11;
Matches 115; Conservative 69; Mismatches 161; Indels 90; Gaps 21;

QY 23 TQGNKV---VLKGG---DTVELTCTASOKKSIQPHMKNSNQIKILNGSFLTKGPSKL 76
DB 4 TTGLKTPSVLKKKTYTQSSQLTTTSE-----VQSSKI 38
QY 77 NDRADRSRLMDQGNFPLIKNLKIEDSDTYICVEVDQKEVQLLVFG--LTANSPTLL 134
DB 39 YCEVVRGSEVWIKI-----IPCKGDKVAPTVILNQSSEELTSRFAFVLCIIDFH-- 91
QY 135 QGQSLLT-----LESPPGSSPVQCRSPRGKN--IQGKTLVSQLELDSGWTCT 185
DB 92 -PESITVSWLKDQGHMESGFVTSPT--C---GVNGTSPATSLTYPAEFTPKRYTQ 144
QY 186 VLQNOKKVEFKIDI---VPCPAPEPKSCDKHTCPPELLGSPVFLPPPKDTLMSRT 241
DB 145 V--SHQGTQSRNITGSOVPC-----SCN-----DVIKLLPPIEGVLT 185
QY 242 PEVTCVVDVSHEDPEYKFNWYVDGVEVNAKTKPREQYNSTYRVSVTLVHODMNG 301
DB 186 VTLNLCI---VSNAIPYGVNSW-----TQEQKSLKSEIAVOPGEDADSVISTQAMLSG 239
QY 302 KEYKCKVSNKALPAPIEKTISKAKQD-REPOV-YTLPPSRDELTKNOVSLTCLVKGFP 359
DB 240 AEFYCVVNHQDLPPTLRASIHKEEVADLREPEVSILLSPAEDVSAQRFSLTCLVKGFS 299
QY 360 SDIAVESNGQPEN--NYKTPPYLDS-DGFFLYSKLTVDKSRWQGNVSCVMHEA 416
DB 300 REIFVKMTVNDKSNVPGNYKTEVMAENDNSIFYISLISIAEEMWAGSCVVGHEA 359
QY 417 LHNHYTOKSLSPG 431
DB 360 IPLKIINRTVKNSSG 374

RESULT 55
ALC1_GORGO
ID ALC1_GORGO STANDARD; PRT; 353 AA.
AC P20758;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig alpha-1 chain C region.
GN IGHA1.
OS Gorilla gorilla gorilla (Lowland gorilla).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Gorilla.
OX NCBI_TaxID=9595;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RX MEDLINE=89386006; PubMed=2506527;
RA Kawamura S., Omoto K., Ueda S.;
RT "Nucleotide sequence of the gorilla immunoglobulin alpha 1 gene.";
RL Nucleic Acids Res. 17:6732-6732(1989).
CC -1- FUNCTION: Ig alpha is the major immunoglobulin class in body
CC secretions. It may serve both to defend against local infection
CC and to prevent access of foreign antigens to the general
CC immunologic system.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15045; CAA3147.1; ALT_INIT.
CC HSSP; P01810; 2F8J.
CC InterPro; IPR007110; Ig-1-like.
CC InterPro; IPR003597; Ig_c1.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00407; IGc1; 2.
CC PROSITE; PS00835; IG_LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein. Repeat.
FT NON_TER 1 1
FT DOMAIN 6 98
FT DOMAIN 125 220
FT DOMAIN 228 330
FT DISULFID 14 14
FT DISULFID 26 85
FT DISULFID 77 101
FT DISULFID 122 122
FT DISULFID 123 180
FT DISULFID 147 204
FT DISULFID 182 182
FT DISULFID 192 192
FT DISULFID 250 313
FT DISULFID 352 352
FT CARBOHYD 144 144
FT CARBOHYD 340 340
FT SEQUENCE 353 AA; 37755 MW; 4820EBDB02AC7514 CRC64;
Query Match 10.3%; Score 279; DB 1; Length 353;
Best Local Similarity 28.0%; Pred. No. 6.6e-11;
Matches 96; Conservative 41; Mismatches 126; Indels 80; Gaps 16;
OY 127 ANSDTHLLOGSLTTLSPSSPSVOCSPRGRKNIQGGKTLVSQLELDQSGTWCTV 186
DB 59 ASGDLYTSSQ---LTL-----PATQC--PDGKSV-----TGHV 87
OY 187 LONOKVFEKIDIVPC-----PAPPEKSCDKHTTCPELLGSPVFLPPKP 232
DB 88 --NHYTNPSQDVTVPORVSPPTPSPSPPTPSPPCCH-----PRLSLHRPAL 134
OY 233 KDTLMIISREVTGVVVDVSHEDPEYKFMVYDVGEVHNAKTPREQYNSTYRVSVLT 292
DB 135 EL-LIGSRLNLTITLTLGL-RDASGVTFWTWPSGK--SAVGPERRDLGCGYSVSSVLP 190
OY 293 VLAQDMLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPQVYTLPPSRDELTKNQ-VSLT 351

```

```

DB 191 GCAEPNMHNGKFTCTAAVPSKPTLTATLSKS-GNMFREPVHLLPPSEELALNELVLT 249
OY 352 CLVKGFPSPDIAYWESNGC--PENNYKTPPVLD-SDG--SFFLYSLTLYDKRRMOGN 406
DB 250 CLARGFSPKVDLVNWLQSGDELPREKYLTVNARSQEPGQTTTFAVSTILRVAAEDWKGD 309
OY 407 VFSGSVNHEALNHYTKSLSLSPG-----LQLDETC 438
DB 310 TFSQVCHGHALPLAFQTKTIDRLAGKPTNVNVSVMAEVDGTC 352
RESULT 56
ALCI HUMAN STANDARD; PRT; 353 AA.
AC P01876;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ig alpha-1 chain C region.
GN IGHAI1
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84130179; PubMed=6421489;
RA Flanagan J.G., Lefranc M.-P., Rabbitts T.H.;
RT "Mechanisms of divergence and convergence of the human immunoglobulin
RT alpha 1 and alpha 2 constant region gene sequences.";
RL Cell 36:681-688(1984).
RN [2]
RP SEQUENCE (MYELOMA PROTEIN BUR), AND DISULFIDE BONDS.
RX MEDLINE=79151016; PubMed=107164;
RA Putnam F.W., Liu Y.-S.V., Low T.-L.K.;
RT "Primary structure of a human IgM1 immunoglobulin. IV. Stereotyped
RT IgM1 protease, digestion, Fab and Fc fragments, and the complete
RT amino acid sequence of the alpha 1 heavy chain.";
RL J. Biol. Chem. 254:2865-2874(1979).
RN [3]
RP SEQUENCE (MYELOMA PROTEIN TRO).
RX MEDLINE=76023781; PubMed=809311;
RA Kratzin H., Altevogt P., Ruban E., Kortz A., Starosciak K.;
RA Hilschmann N.;
RT "The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro.),
RT II. The amino acid sequence of the H-chain, alpha-type, subgroup III;
RT structure of the complete IgA-molecule.";
RL Hoppe-Seyler's Z. Physiol. Chem. 356:1337-1342(1975).
RN [4]
RP DISULFIDE BONDS.
RX MEDLINE=80114124; PubMed=393607;
RA Yang C.-Y., Kratzin H., Gotz H., Hilschmann N.;
RT "Rule of antibody structure. Primary structure of a human monoclonal
RT IgA-immunoglobulin (myeloma protein Tro). VII. Purification and
RT characterization of the disulfide bridges.";
RL Hoppe-Seyler's Z. Physiol. Chem. 360:1919-1940(1979).
RN [5]
RP REVIEW.
RX MEDLINE=91054387; PubMed=2241915;
RA Kerr M.A.;
RT "The structure and function of human IgA.";
RL Biochem. J. 271:285-296(1990).
CC -1- FUNCTION: Ig alpha is the major immunoglobulin class in body
CC secretions. It may serve both to defend against local infection
CC and to prevent access of foreign antigens to the general
CC immunologic system.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: J00220; AAC82528.1; ALT_INIT.
 DR PIR: A23360; ALHU.
 DR PDB: 1IGA; 15-JUN-99.
 DR Genew: HGNC:5478; IGH1.
 DR MIM: 146900; .
 DR GO: GO:0003823; F:antigen binding; NAS.
 DR GO: GO:0006955; P:immune response; NAS.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003597; IG cl.
 DR InterPro: IPR003006; IG_MHC.
 DR Pfam: PF00047; Ig; 3.
 DR SMART: SM00407; IgC1; 2.
 DR PROSITE: PS50835; IG_LIKE; 3.
 DR PROSITE: PS00290; IG_MHC; 1.
 KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KM Polymorphism: 3D-structure; Repeat.
 FT NON_TER 1 1
 FT DOMAIN 6 98
 FT DOMAIN 125 220
 FT DOMAIN 228 330
 FT DISULFID 14 14
 FT DISULFID 26 85
 FT DISULFID 77 101
 FT DISULFID 122 122
 FT DISULFID 123 180
 FT DISULFID 147 204
 FT DISULFID 182 182
 FT DISULFID 192 192
 FT DISULFID 250 313
 FT DISULFID 352 352
 FT CARBOHYD 105 105
 FT CARBOHYD 111 111
 FT CARBOHYD 113 113
 FT CARBOHYD 119 119
 FT CARBOHYD 121 121
 FT CARBOHYD 144 144
 FT CARBOHYD 340 340
 FT VARIANT 176 176
 FT CONFLICT 163 165
 FT CONFLICT 176 176
 FT CONFLICT 190 190
 FT CONFLICT 227 227
 FT CONFLICT 231 231
 FT CONFLICT 290 290
 SQ SEQUENCE 353 AA; 37654 MW; EBA11ECB7E85DB21 CRC64;

Query Match 10.3%; Score 277.5; DB 1; Length 353;
 Best Local Similarity 27.2%; Pred. 0.83e-11;
 Matches 104; Conservative 51; Mismatches 153; Indels 75; Gaps 18;

QY 92 FPLIINKLIKIEDSDTYICEVEDQKEVOLVF-----GLT-----ANSDTLILOG 136
 DB 9 FPLSLCSTQPDGNNVIAICLVQGFPPQEPPLSTVMSSESGQVTANRPPSDASDLTYTSS 68
 QY 137 QSLITLTLBSPGSSPSVQCRSPRGKNIQGGKTLISVQLELDQSGTTCVTYLNQKKVERK 196
 DB 69 Q---LTL-----PATQCLA--GKSV-----TCHVKHYTNPSQ-- 95
 QY 197 IDI-VPCPAPBPKSCCKHTTCEPLDGG---PSVFLFPKPKXTLMTSRPEVTCVVVDS 252
 DB 96 -DVTWCPVPSTPTPTSPSTPTPPSPSCCHPRSLTLRPALBD-LTLGSSANLTCYLTGL- 152
 QY 253 HEDPEVKFMWYDGVENNAKTKPREQYNSTYRVSVLTVLHQDWLNKKEYCKKVSNKA 312
 DB 153 RDAAGVTFMTWTPSSGK--SAVGQPPRPDLGCGSVSVSLPGCAEPNNHKTFTCTAYPE 210

QY 313 LAPLEIKTSKAGQPREPOVYTLPSRDELTKNO-VSLTCLVKGFYPSDIAVEMESNQ 371
 DB 211 SKTPPLATLSKS-GMTFPEVHLPPSESLNLNELVYTLCLARGSPDQVVRMQSGQ 269
 QY 372 --PENNYKTPPVLV-SDG--SFPLYSKLTVDKSRMOQGNVSCSVMEALHNHYTKSL 426
 DB 270 ELPREKYLTMASROBPSQCTTTFATVSLIRVAABDMKGDPTSCVNGHEALPLAFQKTI 329
 QY 427 SLSPG-----LQLEDFC 438
 DB 330 DRLAGKPTHVNVVMAAEVDGTC 352

RESULT 57
 MUCM_ICTPU
 ID MUCM_ICTPU STANDARD; PRT; 481 AA.
 AC P23735;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ig mu chain C region membrane-bound form.
 OS Ictalurus punctatus (Channel catfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Siluriformes;
 OC Ictaluridae; Ictalurus.
 OK NCBI_TaxId=9998;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90384824; PubMed=2119496;
 RA Wilson M.R., Marcuz A., van Ginkel F., Miller N.W., Clem L.W.,
 BA Middleton D., Warr G.W.;
 RT "The immunoglobulin M heavy chain constant region gene of the channel
 catfish, *Ictalurus punctatus*: an unusual mRNA splice pattern produces
 RT the membrane form of the molecule.";
 RL Nucleic Acids Res. 18:5227-5233(1990).
 CC -I- MISCELLANEOUS: During differentiation, B lymphocytes switch from
 CC expression of membrane-bound Igm to secretion of Igm. The mu
 CC chains of membrane and secreted Igm differ in their C-terminal
 CC segments.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 CC EMBL: X52617; -, NOT_ANNOTATED_CDS.
 CC HSSP: P01857; IFC1.
 CC InterPro: IPR007110; IG-like.
 CC InterPro: IPR003597; IG cl.
 CC InterPro: IPR003006; IG_MHC.
 CC Pfam: PF00047; Ig; 3.
 CC SMART: SM00407; IgC1; 1.
 CC PROSITE: PS50835; IG_LIKE; 3.
 CC PROSITE: PS00290; IG_MHC; 2.
 KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KM Transmembrane.
 FT NON_TER 1 1
 FT DOMAIN 1 101
 FT DOMAIN 102 203
 FT DOMAIN 204 304
 FT DOMAIN 305 434
 FT TRANSMEM 459 478
 FT DOMAIN 479 481
 FT DISULFID 14 14
 FT DISULFID 26 85
 FT DISULFID 129 188
 FT DISULFID 233 284
 FT DISULFID 329 391

Query Match
 Best Local Similarity 23.1%; Score 276.5; DB 1; Length 481;
 Matches 113; Conservative 72; Mismatches 156; Indels 149; Gaps 23;

```

FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 481 AA; 53337 MW; 668E784E181C71BA CRC64;

Query Match
Best Local Similarity 23.1%; Score 276.5; DB 1; Length 481;
Matches 113; Conservative 72; Mismatches 156; Indels 149; Gaps 23;

37 VELTCTASOKS---IQFMKNSNOIKILGNSGSLTKGPKSLNDRSRSLMDQGNP 93
22 VTLGVTRDLASADGLSFIWKDAS-----GSLVT-----DIVGYR 56
94 LI-----IKNLIEDSD-----TYICEVE-----DQKEVQLLVFGLT 126
57 AVQATGTYTSVSHVVKASDMNGNKKFTCEVNGKLGSKDASLQKEVERELHASLT---LT 113
127 ANSDTHLQGSGLTTLTLESPGSSPSVQCRSPRG-----KNI---QGG 166
114 TPTQIEINGRTATPVCILATP-----FSKSHTPKMLEKTIDISNKVENIVSQNG 164
167 KTLVSQLELDQSGTWT-----CTVLQNKQKVEFKIDIVPCAPAPKSCDKHTCPBL 219
165 NFTAISVLELSAS-EMTSTSTSPKCEFOQKHNHVFKEASVAPGDTKQDQ-----VK 214
220 LGSPVFLFPKPKKOTLMISRTPEYTCVVDVSHDEPKV-FNMYVDVGEVHNKTKRRE 278
215 ITGPT-----DILIKRAGQEC---RAEDTGFPSIKWILGNREISSLSNL--- 259
279 EQYNSTYRVSVLT-VLHODMNGKRYKCVSNKALPAPIEK-TISKAGOPREPOVYTL 336
260 ----SSKTTVSIQTHIGPEEMNGTEPFICEVHEHFTQOYEKVTYKRENGNDEPKVYL 315
337 PPSRDELTKNOVSLTCLYKGFPSDIAYEWESN-----GQPNNTKTPPVLDSDGS 388
316 AP--PESSGESVTLTCYVDPFYKREVAVSWLVNDQVEEVGYEQN---TTAVIDRNLT 369
369 PFLYKLTIVDKSRMOQGVNFGSVNHEALHNHTYQKSLISLSGLDLDTCAADGELDG 448
370 FSVSYQLIITKADNMNGSVFCLVYHESIKDCVHRHSISA-----KDSKTPTLVN 420
449 LWTDPPPRAS 458
421 LTLTNPQSCS 430

```

RESULT 58
 ALC_MOUSE ID ALC_MOUSE STANDARD; PRT; 344 AA.
 AC POL1878;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ig alpha chain C region.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A. (MYELOMAS ABE48 AND J558).
 RX MEDLINE=81261947; PubMed=6790349;
 RA Auffray C., Neggeotte R., Sikorav J.-L., Heidmann O., Rougeon F.;
 RT "Mouse immunoglobulin A: nucleotide sequence of the structural gene
 for the alpha heavy chain derived from cloned cDNAs.";
 RL Gene 13:365-374 (1981).
 [2]
 RP SEQUENCE OF 1-213 (MOPC 47A).
 RX MEDLINE=80049769; PubMed=115869;
 RA Robinson E.A., Appella E.;
 RT "Amino acid sequence of a mouse myeloma immunoglobulin heavy chain

RT (MOPC 47 A) with a 100-residue deletion.";
 RL J. Biol. Chem. 254:11418-11430 (1979).
 RN [3]
 RP SEQUENCE OF 1-254 AND 291-344 (M511).
 RX MEDLINE=81054880; PubMed=6776528;
 RA Robinson E.A., Appella E.;
 RT "Complete amino acid sequence of a mouse immunoglobulin alpha chain
 (MOPC 511)."
 RL Proc. Natl. Acad. Sci. U.S.A. 77:4909-4913 (1980).
 CC -I- FUNCTION: Ig alpha is the major immunoglobulin class in body
 CC secretions. It may serve both to defend against local infection
 CC and to prevent access of foreign antigens to the general
 CC immunologic system.
 CC -I- MISCELLANEOUS: The final C-region domain is deleted from Ref.2
 CC chain. It was isolated from a myeloma protein that contains 1
 CC light and 1 heavy chain per molecule, linked by a disulfide bond.
 CC In contrast, normal mouse IgA molecules contain 2 light and 2
 CC heavy chains and lack a light-heavy chain disulfide bond.
 CC -I- MISCELLANEOUS: M511 chain was isolated from a myeloma protein that
 CC binds phosphorylcholine.
 CC -I- MISCELLANEOUS: M511 sequence was compared with that of mouse MOPC
 CC 47A, and a genetic mechanism for the deletion of the CH3 domain of
 CC the mutant chain is proposed.
 CC -I- SIMILARITY: Contains 3 immunoglobulin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D11468; BAA02026.1; ALT_INIT.
 DR PIR: A91479; AHMS.
 DR HSSP: P01810; ZFBJ.
 DR InterPro: IPR007110; IG_1like.
 DR InterPro: IPR003597; IG_C1.
 DR InterPro: IPR003006; IG_MHC.
 DR Pfam: PF00047; Ig; 3.
 DR SMART: SM00407; IgC1; 2.
 DR PROSITE: PS50835; IG_1like; 3.
 DR PROSITE: PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
 FT NON TER 1
 FT DOMAIN 6
 FT DOMAIN 116
 FT DOMAIN 219
 FT DISULFID 26
 FT DISULFID 76
 FT DISULFID 114
 FT DISULFID 138
 FT CARBOHYD 38
 FT CARBOHYD 101
 FT CARBOHYD 101
 FT CARBOHYD 331
 FT CARBOHYD 331
 FT VARIANT 2
 FT VARIANT 18
 FT VARIANT 67
 FT VARIANT 73
 FT VARIANT 112
 FT VARIANT 135
 FT VARIANT 141
 FT VARIANT 168
 FT VARIANT 212
 FT VARIANT 235
 FT VARIANT 255
 FT VARIANT 295
 FT VARIANT 301
 FT VARIANT 329
 FT VARIANT 331
 FT VARIANT 331
 SQ SEQUENCE 344 AA; 36676 MW; 3694CF99B19A98 CRC64;

Query Match
 9.8%; Score 264.5; DB 1; Length 344;

Best Local Similarity 26.7%; Pred. No. 5,4e-10;
Matches 85; Conservative 57; Mismatches 117; Indels 59; Gaps 15;

QY 136 GGGTLTLESPPG-----SSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTW 182
DB 45 GKQIT-TVNPPPLASGGRYMSNOULTPAVECC-PEGESVK-----83
QY 183 TCTVLNOKKVEKIDIVPCAPPEKSCDHTHCPELGGSPVFLPPPKKTKMTLSRTP 242
DB 84 -CSVQSDSNPVQ-ELD-VNCSGPTPPPTITPSCQ-----PSLSDRPALDD-LLLGSDA 134
QY 243 EYTCVVVDVSHEDPEVKFMYVGVGVHNAKTKRPREQVNST---YRVSVLTVLTHQDM 299
DB 135 SITCTINGLANRPGAV-FTW-----EPSTCKDAVQKKAVNCGCGCVSVSVLPGCAERN 188
QY 300 NGKEYICKKSNKLPAPRIETIKAKQPREPOVYTLPPSRDELTKNQ-VSLTCLVKGF 358
DB 189 SGASFCYCTVTHPE-SCITLTGTLAKVTVNTPRPQVHLPPPEBELALNELSLTCLVRATN 247
QY 359 PSDIAYEVESNQ--PENNYKTPPVLDSG--SFPLVSKLVDSKRNQGVNFGCSVW 413
DB 248 PKEVLVRLHGNELSPESYLVEPLKEPGEATYLVLSVLRSAETWKQGGQYSCWVG 307
QY 414 HEALHNHYTQKSLSLSPG 431
DB 308 HEALPNFTOKTIDRLSG 325

RESULT 59
SHS1_MOUSE STANDARD; PRT: 513 AA.
ID SHS1_MOUSE 089757; 035924; 088555; 088556; P9796; Q8R559; Q9QX57;
AC 09WTM4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein-tyrosine phosphatase non-receptor type substrate 1 precursor
DE (SHP substrate-1) (SHPS-1) (Inhibitory receptor SHPS-1) (Signal-
DE regulatory protein alpha-1) (SIRP-alpha1) (MSIRP-alpha1) (MyD-1
DE motif) (Blt) (p84).
DE motifa) (Blt) (p84).
GN PTPNS1 OR SHPS1 OR SIRP OR MYD1 OR BIT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX SHS1_MOUSE
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=97233399; PubMed=9070220;
RA Yamao T., Matosaki T., Amano K., Matsuda Y., Takahashi N., Ochi F.,
RA Fujioaka Y., Kasuga M.;
RT "Mouse and human SHPS-1: molecular cloning of cDNAs and chromosomal
RT localization of genes.";
RL Biochem. Biophys. Res. Commun. 231:61-67(1997).
RN [12]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANTS ALA-29; ARG-67; ARG-91;
RP THR-96; SER-128; PRO-194 AND ASN-224.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=97230468; PubMed=9073522;
RA Ohnishi H., Kubota M., Sano S.-I.;
RT "BIT (Blt) maps to mouse chromosome 2.";
RL Genomics 40:504-506(1997).
RN [13]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), SEQUENCE OF 32-53 AND 422-433
RP (ISOFORM 2), VARIANTS ALA-29; ARG-91; THR-96 AND SER-128,
RP N-GLYCOSYLATION, AND TISSUE SPECIFICITY.
RC STRAIN=BALB/c; TISSUE=Brain, and Cerebellum;
RX MEDLINE=98012243; PubMed=9348339;
RA Comu S., Meng W., Olinisky S., Ishwad P., Mi Z., Hempel J., Watkins S.,
RA Lagenaur C.F., Narayanan V.;
RT "The murine P84 neural adhesion molecule is SHPS-1, a member of the
RT phosphatase-binding protein family.";

RL J. Neurosci. 17:8702-8710(1997).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), VARIANTS ALA-29; ARG-67;
RP PRO-74; ALA-83; 86-TYR-VAL-87; ILE-90; ARG-91; THR-96; ALA-114;
RP ILE-118; SER-128; PRO-194; ASN-224; PHE-351 AND ASP-365,
RP N-GLYCOSYLATION, PHOSPHORYLATION ON TYROSINE RESIDUES, INTERACTION
RP WITH PTPN6, AND TISSUE SPECIFICITY.
RC STRAIN=C57BL/6; TISSUE=Fetal thymus;
RX MEDLINE=98380500; PubMed=9712903;
RA Veilleux A., Thibaudan E., Latour S.;
RT "High expression of inhibitory receptor SHPS-1 and its association
RT with protein tyrosine phosphatase SHP-1 in macrophages.";
RL J. Biol. Chem. 273:22719-22728(1998).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3), VARIANTS ALA-29; ARG-67;
RP PRO-74; ALA-83; 86-TYR-VAL-87; ILE-90; ARG-91; THR-96; ALA-114;
RP ILE-118; SER-128; PRO-194; ASN-224; PHE-351 AND ASP-365, AND
RP N-GLYCOSYLATION.
RC STRAIN=129/SvJ, and C57BL/6; TISSUE=Brain, and Liver;
RX MEDLINE=20053880; PubMed=10585853;
RA Sano S.-I., Ohnishi H., Kubota M.;
RT "Gene structure of mouse Bit/SHPS-1.";
RL Biochem. J. 344:667-675(1999).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANTS ALA-29; ARG-67;
RP PRO-74; ALA-83; 86-TYR-VAL-87; ILE-90; ARG-91; THR-96; ALA-114;
RP ILE-118; SER-128; PRO-194; ASN-224; PHE-351 AND ASP-365.
RC STRAIN=ILS, and ISS;
RX MEDLINE=21363810; PubMed=11471062;
RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Caniff J.,
RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikela J.M.;
RT "High-throughput sequence identification of gene coding variants
RT within alcohol-related QTLs.";
RL Mamm. Genome 12:657-663(2001).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT PRO-194.
RC TISSUE=Placenta;
RA Wang H., Chen Z., Ullrich A.;
RT "Epidermal growth factor-induced association of SHP2 with mouse SIRP-
RT alpha1.";
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
RN [8]
RP FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX MEDLINE=90152134; PubMed=2303162;
RA Chuang W., Lagenaur C.F.;
RT "Central nervous system antigen P84 can serve as a substrate for
RT neurite outgrowth.";
RL Dev. Biol. 137:219-232(1990).
RN [9]
RP N-GLYCOSYLATION, PHOSPHORYLATION BY JAK2 IN RESPONSE TO GROWTH
RP HORMONE, AND INTERACTIONS WITH JAK2 AND PTPN11.
RX MEDLINE=98175985; PubMed=9507023;
RA Stofega M.R., Wang H., Ullrich A., Carter-Su C.;
RT "Growth hormone regulation of SIRP and SHP-2 tyrosyl phosphorylation
RT and association.";
RL J. Biol. Chem. 273:7112-7117(1998).
RN [10]
RP INTERACTION WITH CD47, AND TISSUE SPECIFICITY.
RX MEDLINE=9901586; PubMed=9872987;
RA Jiang P., Lagenaur C.F., Narayanan V.;
RT "Integrin-associated protein is a ligand for the P84 neural adhesion
RT molecule.";
RL J. Biol. Chem. 274:559-562(1999).
RN [11]
RP INTERACTION WITH FcR.
RX MEDLINE=20130295; PubMed=10662797;
RA Greban H.D., Dale B.M., Porter J.W., Chang P.W., Vines C.M.,
RA Lowell C.A., Lagenaur C.F., Willman C.L.;
RT "Negative regulation of phagocytosis in murine macrophages by the Src
RT kinase family member, Fgr.";
RL J. Exp. Med. 191:515-526(2000).
CC -1- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts
as docking protein and induces translocation of PTPN6,

CC PRPN1 and other binding partners from the cytosol to the
 CC plasma membrane. Supports attachment of cerebellar neurons, neurite
 CC outgrowth and glial cell attachment. May play a key role in
 CC intracellular signaling during synaptogenesis and in synaptic
 CC function. Involved in the negative regulation of receptor tyrosine
 CC kinase-coupled cellular responses induced by cell adhesion, growth
 CC factors or insulin. Mediates negative regulation of phagocytosis,
 CC mast cell activation and dendritic cell activation. CD47 binding
 CC prevents maturation of immature dendritic cells and inhibits
 CC cytokine production by mature dendritic cells (By similarity).
 CC - SUBUNIT: Binds PRPN1 when tyrosine-phosphorylated, except in
 CC macrophages, where it primarily binds PRPN6. Binds GRB2 *in vitro*.
 CC Binds FGR. Binds JAK2 irrespective of its phosphorylation status
 CC and forms a stable complex. Binds SCAP1 and/or SCAP2. The
 CC resulting complex recruits FVB. Binds PTK2B (By similarity).
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1; Synonyms=a;
 CC IsoId=P97797-1; Sequence=Displayed;
 CC Name=2; Synonyms=a; Large;
 CC IsoId=P97797-2; Sequence=VSP_007032;
 CC Name=3; Synonyms=b, Small;
 CC IsoId=P97797-3; Sequence=VSP_007031;
 CC - TISSUE SPECIFICITY: Highly expressed in cerebral cortex, brain,
 CC spinal cord, cerebellum and spleen, and at much lower levels in
 CC kidney, thymus, heart, lung and liver. Within the cerebellum,
 CC highly expressed throughout the molecular layer, and in synaptic
 CC glomeruli in the granule cell layer. Detected in neurons of the
 CC hippocampus and dentate gyrus, and in olfactory bulb. Not detected
 CC in Purkinje cells. Highly expressed in the plexiform layers, optic
 CC fiber layer and the outer segments of the photoreceptor layer in
 CC the retina. Highly expressed in macrophages. Isoform 3 is detected
 CC at very low levels in all tissues tested.
 CC - DEVELOPMENTAL STAGE: Highly expressed in the CNS of embryos from
 CC day 7 to 17.
 CC - PTM: N-glycosylated.
 CC - PTM: Phosphorylated on tyrosine residues.
 CC - SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.
 CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 CC DR EMBL, D87967; BAA13520.1; -;
 CC DR EMBL, D87968; BAA13521.1; -;
 CC DR EMBL, D87969; BAA20376.1; -;
 CC DR EMBL, U89694; AAB92591.1; -;
 CC DR EMBL, AF072543; AAC24886.1; -;
 CC DR EMBL, AF072544; AAC24887.1; -;
 CC DR EMBL, AB024507; BAA89290.1; -;
 CC DR EMBL, AB024500; BAA89290.1; JOINED.
 CC DR EMBL, AB024501; BAA89290.1; JOINED.
 CC DR EMBL, AB024502; BAA89290.1; JOINED.
 CC DR EMBL, AB024503; BAA89290.1; JOINED.
 CC DR EMBL, AB024504; BAA89290.1; JOINED.
 CC DR EMBL, AB024505; BAA89290.1; JOINED.
 CC DR EMBL, AB024506; BAA89290.1; JOINED.
 CC DR EMBL, AB018194; BAA76555.1; -;
 CC DR EMBL, AB024507; BAA89289.1; -;
 CC DR EMBL, AB024500; BAA89289.1; JOINED.
 CC DR EMBL, AB024501; BAA89289.1; JOINED.
 CC DR EMBL, AB024502; BAA89289.1; JOINED.
 CC DR EMBL, AB024503; BAA89289.1; JOINED.
 CC DR EMBL, AB024504; BAA89289.1; JOINED.
 CC DR EMBL, AB024505; BAA89289.1; JOINED.
 CC DR EMBL, AB024506; BAA89289.1; JOINED.

DR EMBL, AF332079; AAK56107.1; -;
 DR EMBL, AF332080; AAK56108.1; -;
 DR EMBL, Y10349; CAA71375.1; -;
 DR PIR, J52589; J52589.
 DR MGI, MGI:108563; Pcpn1.
 DR GO, GO:0008580; F:cytoskeletal regulator activity; IMP.
 DR GO, GO:0045309; F:phosphoprotein amino acid binding; IPI.
 DR GO, GO:0005515; F:protein binding; IPI.
 DR GO, GO:0007015; P:actin filament organization; IMP.
 DR GO, GO:0006928; P:cell motility; IMP.
 DR GO, GO:0007160; P:cell-matrix adhesion; IMP.
 Query Match 7.4%; Score 199; DB 1; Length 513;
 Best Local Similarity 19.8%; Pred. No. 1.4e-05;
 Matches 100; Conservative 63; Mismatches 166; Indels 176; Gaps 17;
 QY 23 TQGNKVVYLGKKGDVELTCTASQKKSIOFHKNSNQIKLQSGFLTKGSKLDRADS 82
 DB 37 TQPEKSVSVAGDSTVNTCTLT-----SLPVGPIK----- 67
 QY 83 RSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOLVGLTANSPTHLQGSLLTLT 142
 DB 68 ----WYRG-----VGQSLILISFTGE--HF----- 87
 QY 143 LESPFGSSPSVQCRSPRKNIOGKTLVSQLDLQDSGTWTCTVLQNKVEFKDIVPC 202
 DB 88 -----PRTVNVSDATKRNMMDFSRISNVTPEDAGTYCVKFO----- 125
 QY 203 PAPERKSCDKHTTPELLGGSVLF---PK---PKDTLMISRTPEVTCVVVDVSH-- 253
 DB 126 ----KGPSEPDTEIQSGGTEVYVLAKPSPPEVSGPADROIPOQKNFTC---KSHGF 176
 QY 254 EDPVKEFWYVDGVEVNAKTPREQVNSYRVSVLTVLHQDLNKGEXKCVSKAL 313
 DB 177 SPRNITLKFWDGGLHLELTTPVSGKNVSYNISVYRV---LNSMDVSKYICVA 232
 QY 314 PAPIEKTISKAGQPREPOVYTLPPS-----RDELTKNOVSLTLYKGFYSDIAVENES 368
 DB 233 HITLDR--SPRGIANLSNFRVSPVTKVQGSPTSMNQVVLTRAEERFYEDQLWLE 290
 QY 369 NGQENNNYKTPPLV--DSQGSFPLYSKLYDKSRMOQGNVFSQVWHE---ALHNHTQ 423
 DB 291 NGVNSRN--DTPKNLTAKTQDGTNYVTSLFLVNSSAHEDVDVFTQVKHDDQPAITRNHTV 348
 QY 424 KSLSLSPQLQDERCAEQDELGLMTTDPRASALPAPETGALPDPOTASALPDPEA 483
 DB 349 LGLHNS-----SDQGS 359
 QY 484 ASALPALAVISFLGLGLGVCYL 508
 DB 360 MQTFPGNNATHNMNVFIGVAVCAL 384
 RESULT 60
 SHS1 RAT STANDARD; PRT; 509 AA.
 AC P97710; O08951; O70426; O90W15;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Protein-tyrosine phosphatase non-receptor type substrate 1 precursor
 DE (SHP substrate-1) (SHPs-1) (Inhibitory receptor SHPS-1) (Signal-
 DE regulatory protein alpha-1) (SHP-alpha-1) (Brain Ig-like molecule
 DE with tyrosine-based activation motif) (Blt) (Macrophage fusion
 DE receptor) (Macrophage membrane protein MFP150).
 GN PTENSI OR SHPS1 OR SIRP OR BIR OR MFR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 42-60; 68-91; 128-137; 150-158;
 RP 174-189; 192-202; 204-212; 218-237; 259-270; 279-282; 405-415 AND

RP 446-453, N-GLYCOSYLATION, PHOSPHORYLATION ON TYROSINE RESIDUES, AND
 RP INTERACTIONS WITH PTPN6 AND PTPN11.
 RC TISSUE=Fetal fibroblast;
 RX MEDLINE=9709667; PubMed=9434344;
 RA Fujioke Y., Matczak T., Noguchi T., Iwamatsu A., Yamao T.,
 RA Takahashi N., Tauda M., Takada T., Kaunga M.;
 RT "A novel membrane glycoprotein, SHS-1, that binds the SH2 domain-
 RT containing protein tyrosine phosphatase SHP-2 in response to mitogens
 RT and cell adhesion.";
 RL Mol. Cell. Biol. 16:6887-6899(1996).
 RN [2]
 RP SEQUENCE FROM N.A., SEQUENCE OF 32-48 AND 446-453, FUNCTION, AND
 RP PHOSPHORYLATION ON TYROSINE RESIDUES.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=97415431; PubMed=9271230;
 RA Sano S.-I., Ohnishi H., Omori A., Hasegawa J., Kubota M.;
 RT "B17, an immune antigen receptor-like molecule in the brain.";
 RL FEBS Lett. 411:327-334(1997).
 RN [3]
 RP SEQUENCE FROM N.A., SEQUENCE OF 99-107, 128-149, 192-217, 405-417;
 RP 419-429; 446-467 AND 496-506, N-GLYCOSYLATION, AND TISSUE SPECIFICITY.
 RC STRAIN=Fischer 344; TISSUE=Spleen;
 RX MEDLINE=94449911; PubMed=9774638;
 RA Seginario C., Sterling H., Beckers C., Kobayashi R., Solimena M.,
 RA Ullrich E., Vignery A.;
 RT "M μ , a putative receptor mediating the fusion of macrophages.";
 RL Mol. Cell. Biol. 18:6213-6223(1998).
 RN [4]
 RP SEQUENCE OF 1-419 FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN=MAG/Rij; TISSUE=Alveolar macrophage;
 RX MEDLINE=98375871; PubMed=9712053;
 RA Adams S., van der Laan L.J.W., Vernon-Wilson E.,
 RA Remacle de Lavielette C., Doepf E.A., DiKlerka C.D., Simmons D.L.,
 RA van den Berg T.K.;
 RT "Signal-regulatory protein is selectively expressed by myeloid and
 RT neuronal cells.";
 RL J. Immunol. 161:1853-1859(1998).
 RN [5]
 RP PHOSPHORYLATION IN RESPONSE TO EGF, AND INTERACTION WITH PTPN11.
 RX MEDLINE=9800885; PubMed=9344856;
 RC Ochi F., Matczak T., Noguchi T., Fujioke Y., Yamao T., Takada T.,
 RA Tauda M., Takada H., Fukunaga K., Okabayashi Y., Kaunga M.;
 RT "Epidermal growth factor stimulates the tyrosine phosphorylation of
 RT SHS-1 and association of SHS-1 with SHP-2, a SH2 domain-containing
 RT protein tyrosine phosphatase.";
 RL Biochem. Res. Commun. 239:483-487(1997).
 RN [6]
 RP PHOSPHORYLATION OF TYR-477 AND TYR-501, AND MUTAGENESIS OF TYR-436;
 RP TYR-460; TYR-477 AND TYR-501.
 RX MEDLINE=98049923; PubMed=9535915;
 RA Takada T., Matczak T., Takada H., Fukunaga K., Noguchi T.,
 RA Fujioke Y., Okazaki I., Tauda M., Yamao T., Ochi F., Kaunga M.;
 RT "Role of the complex formation of SHS-1 with SHP-2 in
 RT insulin-stimulated mitogen-activated protein kinase activation.";
 RL J. Biol. Chem. 273:9234-9242(1998).
 RN [7]
 RP FUNCTION. Immunoglobulin-like cell surface receptor for CD47. Access
 RP as docking protein and induces translocation of PIPN6,
 RP PIPN1 and other binding partners from the cytosol to the
 RP plasma membrane. Supports adhesion of cerebellar neurons, neurite
 RP outgrowth and glial cell attachment. May play a key role in
 RP intracellular signaling during synaptogenesis and in synaptic
 RP function. Involved in the negative regulation of receptor tyrosine
 RP kinase-coupled cellular responses induced by cell adhesion, growth
 RP factors or insulin. Mediates negative regulation of phagocytosis,
 RP mast cell activation and dendritic cell activation. CD47 binding
 RP prevents maturation of immature dendritic cells and inhibits
 RP cytokine production by mature dendritic cells. May play a role in
 RP the release of nitric oxide by macrophages (By similarity).
 RP -1 SUBUNIT: Binds PTPN11 when tyrosine-phosphorylated, except in
 RP macrophages, where it primarily binds PTPN6. Binds GNB2 in vitro.
 RP Binds PDK. Binds JAK2 irrespective of its phosphorylation status
 RP and forms a stable complex. Binds SCAPI and/or SCAPI2. The
 RP resulting complex recruits Fyb. Binds PKC ζ (By similarity).

CC	-1	SUBCELLULAR LOCATION: Type I membrane protein.	
CC	-1	TISSUE SPECIFICITY: Highly expressed in brain, spleen, lung, liver and kidney. Detected at lower levels in heart. Highly expressed in alveolar and peritoneal macrophages, and at lower levels in dendritic cells.	
CC	-1	PTM: N-glycosylated.	
CC	-1	PTM: Phosphorylated on tyrosine residues in response to insulin, cell adhesion or epidermal growth factors. Dephosphorylated by PTPN11.	
CC	-1	SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.	
CC	-1	SIMILARITY: Contains 1 immunoglobulin-like V-type domain.	
CC	-1	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).	
CC		EMBL; D85183; BAA12734.1; -.	
DR		EMBL; D8468; BAA20368.1; -.	
DR		EMBL; U62328; AAC68478.1; -.	
DR		EMBL; AF055065; AAC18089.1; -.	
DR		HSSP; P01703; 7FAB.	
DR		InterPro; IPR007110; Ig-like.	
DR		InterPro; IPR003597; Ig cl.	
DR		InterPro; IPR003006; Ig_MHC.	
DR		Pfam; PF00047; Ig; 3.	
DR		SMART; SM00407; Tgcl; 2.	
DR		PROSITE; PS50835; Ig_LIKE; 3.	
DR		PROSITE; PS00290; Ig_MHC; 1.	
KW		Repeat; Signal; Transmembrane.	
KW		Glycoprotein; Phosphorylation.	
FT		SIGNAL	1
FT		CHAIN	32
FT			509
FT		PROTEIN-TYROSINE PHOSPHATASE NON-RECEPTOR TYPE SUBSTRATE 1.	
FT		DOMAIN	28
FT		TRANSMEM	374
FT			394
FT		DOMAIN	395
FT			509
FT		DOMAIN	32
FT			138
FT		DOMAIN	150
FT			248
FT		DOMAIN	255
FT			349
FT		DISULFID	55
FT			122
FT		DISULFID	172
FT			229
FT		DISULFID	274
FT			332
FT		SITE	436
FT			439
FT		SITE	446
FT			451
FT		SITE	460
FT			463
FT		SITE	477
FT			480
FT		SITE	501
FT			504
FT		MOD_RES	436
FT			436
FT		MOD_RES	460
FT			460
FT		MOD_RES	477
FT			477
FT		MOD_RES	501
FT			501
FT		CARBOHYD	54
FT			54
FT		CARBOHYD	93
FT			93
FT		CARBOHYD	169
FT			169
FT		CARBOHYD	181
FT			181
FT		CARBOHYD	205
FT			205
FT		CARBOHYD	209
FT			209
FT		CARBOHYD	242
FT			242
FT		CARBOHYD	246
FT			246
FT		CARBOHYD	271
FT			271
FT		CARBOHYD	293
FT			293
FT		CARBOHYD	312
FT			312
FT		CARBOHYD	320
FT			320
FT		CARBOHYD	345
FT			345
FT		MUTAGEN	436
FT			436
FT		Y-F: ABOLISHES TYROSINE PHOSPHORYLATION AND PTPN11 BINDING, WHEN ASSOCIATED WITH F-460 F-477 AND F-501.	


```

CC -1- FUNCTION: Probable immunoglobulin-like cell surface receptor.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9P1W8-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9P1W8-2; Sequence=VSP_007027;
CC Note=No experimental confirmation available;
CC Name=3;
CC IsoId=Q9P1W8-3; Sequence=VSP_007028;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Detected in liver, and at very low levels in
CC brain, heart, lung, pancreas, kidney, placenta and skeletal
CC muscle.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AB042824, BAA95692.1, -.
CC EMBL, AL138804, CAC00474.1, -.
CC EMBL, BC020629, AAH20629.1; ALT_INIT.
CC GeneW, HGNC:15757, SIRPB2.
CC MIM, 605466, -.
CC DR GO, GO:0007267, P:cell-cell signaling; TAS.
CC DR GO, GO:0007242, P:intracellular signaling cascade; TAS.
CC DR GO, GO:0008285, P:negative regulation of cell proliferation; TAS.
CC DR InterPro, IPR007110, Ig_1like.
CC DR InterPro, IPR003597, Ig_c1.
CC DR InterPro, IPR003006, Ig_MHC.
CC DR InterPro, IPR003596, Ig_v.
CC DR Pfam, PF00047, Ig_3.
CC DR SMART, SM00407, IgC1, 2.
CC DR SMART, SM00406, IgV, 1.
CC DR PROSITE, PS00835, IG_LIKE, 3.
CC DR PROSITE, PS00290, IG_MHC; FALSE NEG.
CC KM Repeat; Signal; Transmembrane; Immunoglobulin domain; Glycoprotein;
CC Alternative splicing.
CC FT SIGNAL 1 28 POTENTIAL.
CC FT CHAIN 29 387 SIGNAL-REGULATORY PROTEIN BETA-2.
CC FT DOMAIN 29 360 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 361 383 POTENTIAL.
CC FT DOMAIN 384 387 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 29 137 IG-LIKE V-TYPE.
CC FT DOMAIN 146 245 IG-LIKE C1-TYPE 1.
CC FT DOMAIN 252 340 IG-LIKE C1-TYPE 2.
CC FT DISULFID 53 119 POTENTIAL.
CC FT DISULFID 168 226 POTENTIAL.
CC FT CARBOHYD 221 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT VARSPLIC 1 33 Missing (in isoform 2).
CC FT VARSPLIC 33 /FridaVSP_007027.
CC FT VARSPLIC 144 360 Missing (in isoform 3).
CC FT VARSPLIC 360 /FridaVSP_007028.
CC FT CONFLICT 263 263 A -> V (IN REF. 2).
CC FT CONFLICT 286 286 L -> S (IN REF. 2).
CC SO SEQUENCE 387 AA; 42495 MW; F7F20C9F8680E4B CRC64;
Query Match 7.0%; Score 188.5; DB 1; Length 387;
Best Local Similarity 21.9%; Pred. No. 4.5e-05;
Matches 113; Conservative 56; Mismatches 174; Indels 173; Gaps 20;
QY 6 PFRHLLLVQLMLLPA-----TQGNVVLGKGDVLELCTASQKSIQFMHNSNG 58

```

```

DB 13 PF--LTLTLTLTLTAVAGEBELQMTQPEKLLVYTKATLACTYT----- 56
QY 59 IKILNQGSFLLTKGSPSKLNDRADSRRLMDQGNP--LIINKLIEDSDTYICEVEDQK 115
DB 57 -----SLLPVGP-----VLMFRGVGPRELITY-----NQK 81
QY 116 EKVQVLVRLTANSDTHLQGSLLVLTLESPPGSSPSVQCRRPRGKNIQGKTLVSQLE 175
DB 82 E-----GHPRTVTVSDLTGRNNMDFSRISIT 110
QY 176 LQDSGTWCTVLQ--NOKKVEKIDIVCPAPEPKSCDKHTCPQLGSPV-FLFPPKP 232
DB 111 PADVGTTCVKKRKSPEVVERK-----SGPGTEMLGAKRP 146
QY 233 KDTLMIS--RPEVTCVVVDVSH--EDPEVKFMVYDGVGEVHNAKTKPREEYNSTYR 286
DB 147 SAPVILGPAARTPTPHVTSFTCESHGFSRDTLLKMFKNGNELSPQFVVDPTQSVAYS 206
QY 287 VVSVLTVLHQDPLNGEKYKCKSNKALRA-PIEKT--ISKAGQREPOVYTLPSRDEL 343
DB 207 IRTSTARVVLDPDVVSQVYCEVAHTVLQGDPLRGTMNSEAIRVPEPLEV-TQQPWR--- 262
QY 344 TKNQVSLTCLVKGFPYSDIAVEESNGQENNYKTPPYLDSGSEFPYLSKLYDKSRMQ 403
DB 263 AGNQVAVTCQVKKFPOSILQTLWENGVCQRETSLTLENKQGYTNMTSWPLVNSDQR 322
QY 404 QGNVSCSVMEALHNHYTKSLSPGLQIDETCAEQDGLDGLMTTPDPPASALPAP 463
DB 323 DDVVLTCQVKGHDG-----QLAVSKRLALEVTYHQXD----- 353
QY 464 PTGSALPDPQTASALPDPASALPALAVISFLLG 499
DB 354 -----QSSDATPGR--ASST-TALLLAVLLG 377
RESULT 62
SHS1_BOVIN STANDARD; PRT; 506 AA.
ID SHS1_BOVIN
AC 046631; 046632;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein-tyrosine phosphatase; non-receptor type substrate 1 precursor
DE (SHP substrate-1) (SHPs-1) (inhibitory receptor SHS-1) (Signal-
DE regulatory protein alpha-1) (S1rp-alpha-1) (MyD-1 antigen).
DE PTPNS1 OR SHPS1 OR S1RP OR MYD1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A., AND VARIANTS SER-23; ALA-28; LEU-61; ARG-70;
RP HIS-120; I25-GLN; I27-GLY; I29-HIS; I32-VAL; ASN-145; VAL-155;
RP ASP-203; ARG-261; LEU-302; LEU-316; ARG-337; ASN-367; LEU-422; PHE-429
RP AND GLU-433.
RC STRAIN=Frilesian; TISSUE=peripheral blood;
RX MEDLINE=98143722; PubMed=9485180;
RA Brooke G.P., Parsons K.R., Howard C.J.;
RT "Cloning of two members of the S1RP alpha family of protein tyrosine
RT phosphatase binding proteins in cattle that are expressed on monocytes
RT and a subpopulation of dendritic cells and which mediate binding to
RT CD4 T cells."
RL Eur. J. Immunol. 28:1-11(1998).
CC -1- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts
CC as docking protein and induces translocation of PIP2,
CC PIP3 and other binding partners from the cytosol to the
CC plasma membrane. Supports adhesion of cerebellar neurons, neurite
CC outgrowth and glial cell attachment. May play a key role in
CC intracellular signaling during synaptogenesis and in synaptic
CC function. Involved in the negative regulation of
CC receptor tyrosine kinase-coupled cellular responses induced by

```

CC cell adhesion, growth factors or insulin. Mediates negative
 CC regulation of phagocytosis, mast cell activation and dendritic
 CC cell activation. CD47 binding prevents maturation of immature
 CC dendritic cells and inhibits cytokine production by mature
 CC dendritic cells (By similarity).
 CC -1 SUBUNIT: Binds PTPN11 when tyrosine-phosphorylated, except in
 CC macrophages, where it primarily binds PTPN6. Binds GIB2 *in vitro*.
 CC Binds JAK2 irrespective of its phosphorylation status and forms a
 CC stable complex. Binds SCAPI and/or SCAPI2. The resulting complex
 CC recruits Fyb. Binds Fgr and Ptk2B (By similarity).
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1 TISSUE SPECIFICITY: Highly expressed in spleen macrophages.
 CC -1 DETECTED IN skin dendritic cells.
 CC -1 PTM: Phosphorylated on tyrosine residues (By similarity).
 CC -1 SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.
 CC -1 SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL, Y11045; CAA71942.1; -.
 CC DR EMBL, Y11046; CAA71943.1; -.
 CC DR InterPro: IPR007110; Ig-like.
 CC DR InterPro: IPR003587; Ig_C1.
 CC DR InterPro: IPR003006; Ig_MHC.
 CC DR Pfam; PF00047; Ig; 3.
 CC DR SMART; SM00407; IGc1; 2.
 CC DR PROSITE; PS00835; IG_LIKE; 3.
 CC DR PROSITE; PS00290; IG_MHC; FALSE NEG.
 CC Repeat; Signal; Transmembrane; Immunoglobulin domain; SH3-binding;
 CC Glycoprotein; Phosphorylation; Polymorphism.
 CC KX SIGNAL 1 29
 CC CHAIN 30 506
 CC FT DOMAIN 30 371
 CC FT TRANSMEM 372 392
 CC FT DOMAIN 393 506
 CC FT DOMAIN 30 145
 CC FT DOMAIN 148 248
 CC FT DOMAIN 255 348
 CC FT DISULFID 55 121
 CC FT DISULFID 170 228
 CC FT DISULFID 273 331
 CC FT SITE 432 435
 CC FT SITE 441 446
 CC FT SITE 455 458
 CC FT SITE 472 475
 CC FT SITE 498 501
 CC FT MOD_RES 431 431
 CC FT MOD_RES 455 455
 CC FT MOD_RES 472 472
 CC FT MOD_RES 498 498
 CC FT CARBOHYD 92 92
 CC FT CARBOHYD 167 167
 CC FT CARBOHYD 179 179
 CC FT CARBOHYD 204 204
 CC FT CARBOHYD 210 210
 CC FT CARBOHYD 246 246
 CC FT CARBOHYD 270 270
 CC FT CARBOHYD 292 292
 CC FT CARBOHYD 311 311
 CC FT CARBOHYD 319 319
 CC FT CARBOHYD 344 344
 CC FT CARBOHYD 365 365

FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 23 23 T -> S.
 FT VARIANT 28 28 N -> A.
 FT VARIANT 61 61 S -> L.
 FT VARIANT 70 70 K -> R.
 FT VARIANT 120 120 Y -> H.
 FT VARIANT 125 125 E -> Q.
 FT VARIANT 127 127 R -> G.
 FT VARIANT 129 129 R -> H.
 FT VARIANT 132 132 M -> V.
 FT VARIANT 133 132 S -> V.
 FT VARIANT 145 145 S -> N.
 FT VARIANT 153 153 L -> V.
 FT VARIANT 203 203 N -> D.
 FT VARIANT 261 261 G -> R.
 FT VARIANT 302 302 F -> L.
 FT VARIANT 316 316 F -> L.
 FT VARIANT 337 337 G -> R.
 FT VARIANT 367 367 S -> N.
 FT VARIANT 422 422 Q -> L.
 FT VARIANT 429 429 I -> F.
 FT VARIANT 433 433 D -> E.
 SQ SEQUENCE 506 AA; 55093 MM; 6878310677FCF9CB CRC64;
 Query Match 7.0%; Score 188.5; DB 1; Length 506;
 Best Local Similarity 22.4%; Pred. No. 6.3e-05;
 Matches 71; Conservative 52; Mismatches 139; Indels 55; Gaps 11;
 Oy 158 PRGNITGGK-----TLVSQLELDSGTWTCTVLQGNKK--VERKIDVPPAPAPK 208
 Db 88 PRVTVSDATKRNKMDISIRSNITTPADAGYGVCKFKRKGDMERF-----SGPETH 141
 Oy 209 SCDKTHTCPELLGSPVFLFPKPKDTLMTSRPEVTCVVVDVSH--EDPEVKFMVYDG 266
 Db 142 LTVAKSKSPFLSLSPV-----RATEQVNFCTCHGSPRAISLKWPNHG 188
 Oy 267 VEVNNAKTPREQYNSTYRVSVLYVLDQWLNGKEKCKVSKALPA--PIEKTSKA 324
 Db 189 NELASQTSVDPEDNNVSYSINSTTKVILATGDVHSQVTCVAHTVLQGGPPL----- 241
 Oy 325 KGGREQVYTLPPSRBELT-----KNVSLTCLVGFYPEDIANVESNQPPENNYKT 379
 Db 242 RGTNLSETRVPEPTL-ETIGSPAGNQVNTCCVNFYPPHQLQTLWLENGMSRTEAS 300
 Oy 380 PVLDSGFPLVSKLTVDSRMOQGVFSCVWHEBALHNNHYTQKSLSPGLDDETCA 439
 Db 301 VFVANKGTFRQTSWFLVNSAHRNAVVLTCGVHDG-----QPAVSKNLTLEVSAPQ 354
 Oy 440 EAQDEL----DGLWTT 452
 Db 355 DQDTGQTFPNDNWSWTS 371
 RESULT 63
 LAC1 MOUSE STANDARD; PRT; 105 AA.
 ID LAC1 MOUSE
 AC P01843;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 03, Last sequence update)
 DT 15-MAR-2004 (Rel. 41, Last annotation update)
 DE Ig Lambda-1 chain C region.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=83014953; Pubmed=6812053;
 RX Seising B., Miller J., Wilson R., Scorp U.,
 RT "Evolution of mouse immunoglobulin lambda genes.",
 RL Proc. Natl. Acad. Sci. U.S.A. 79:4681-4685(1982).
 RN [2]
 RP SEQUENCE FROM N.A. (MOPC 315).
 RX MEDLINE=81148806; Pubmed=6259534;

```

RA Boethwell A.L.M., Paekind M., Schwartz R.C., Sonenshein G.E.,
RA Gelfer M.L., Baltimore D.,
RT "Dual expression of lambda genes in the MOPC-315 plasmacytoma.",
RL Nature 290:65-67(1981).
RN
RN (3)
RP SEQUENCE FROM N.A. (S943).
RX MEDLINE=82220143; PubMed=6283385;
RA Boethwell A.L.M., Paekind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.,
RT "Somatic variants of murine immunoglobulin lambda light chains.",
RL Nature 298:380-382(1982).
RN
RN (4)
RP SEQUENCE (MYELOMA PROTEIN MOPC 104E).
RX MEDLINE=71107854; PubMed=5276767;
RA Appella E.,
RT "Amino acid sequences of two mouse immunoglobulin lambda chains.",
RL Proc. Natl. Acad. Sci. U.S.A. 68:590-594(1971).
CC -1- MICELANANOUS: The MOPC 315 cell line produces 2 light chains, 1
CC normal lambda-2 chain and 1 abnormal lambda-1 chain that is
CC missing a large part of the V region. The C region sequence (shown
CC here) appears completely normal.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DR EMBL, J00582; AAA51636.1; -.
DR EMBL, J00587; AAB59672.1; -.
DR PIR, A93922; LIMS.
DR PDB, 1JNH; 06-FEB-02.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003597; IG_CL.
DR InterPro: IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KM Immunoglobulin domain; Immunoglobulin C region; 3D-structure.
KV
KV NON_TER 1 1
FT DOMAIN 1 100 IG-LIKE.
FT DISULFID 27 86
FT DISULFID 104 104 INTERCHAIN (WITH HEAVY CHAIN).
FT CONFLICT 19 20 ET -> TE (IN REF. 4).
FT CONFLICT 56 56 Q -> E (IN REF. 4).
FT CONFLICT 75 82 MISSING (IN REF. 4).
FT CONFLICT 81 82 HS -> SH (IN REF. 4).
FT CONFLICT 85 85 S -> SS (IN REF. 4).
FT CONFLICT 96 96 E -> Q (IN REF. 4).
SQ SEQUENCE 105 AA; 11575 MW; A89F2B09BEFC0A18 CRC64;

Query Match 6.6%; Score 177.5; DB 1; Length 105;
Best Local Similarity 35.9%; Pred. No. 4,4e-05;
Matches 37; Conservative 18; Mismatches 43; Indels 5; Gaps 3;

OY 327 OPR-EFOYTTLPSPSDELTKQVSLTCLVKGYPSPDIAVESNSQP-ENNYKTTTPVLVD 384
DB 1 QPKSSPSYTLPEPSSSELETNKATLVCTITPFPVGVTVDMKVDGTPVQGMETTPSKQ 60
OY 385 SDGSFPLYSKLTVDKSRMOQGVFSCSWHEALHNHYOKSL 427
DB 61 SNNKTMASSYLLTTLTARAWERHSSYSCOVTHE---GHVTKSL 100

RESULT 64
ID _VCAL_RAT STANDARD; PRT; 739 AA.
AC P29534;
DT 01-APR-1993 (Rel. 25, Created)

```

Query Match	Score	DB 1	Length	739
01-Apr-1993 (Rel. 25, Last sequence update)	176.5	DB 1	739	
10-Oct-2003 (Rel. 42, Last annotation update)				
Vascular cell adhesion protein 1 precursor (V-CAM 1).				
NCBI TaxID=10116;				
SEQUENCE FROM N.A.				
TISSUE=Lung;				
MEDLINE=92181437; PubMed=1371918;				
Hession C., Moy P., Tizabi R., Chisholm P., Williams C., Wysk M.,				
Burkly L., Myake K., Kincaid P., Lobb R.,				
"Cloning of murine and rat vascular cell adhesion molecule-1."				
Biochem. Biophys. Res. Commun. 183:163-169(1992).				
-1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION				
IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1				
INTEGRIN VIA4 ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL				
TRANSDUCTION. THE VCAM1/VIA4 INTERACTION MAY PLAY A				
PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE				
EMIGRATION TO SITES OF INFLAMMATION.				
-1- SUBCELLULAR LOCATION: Type I membrane protein.				
-1- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as				
well as on macrophage-like and dendritic cell types in both normal				
and inflamed tissue.				
-1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.				

This SWISS-PROT entry is copyright. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL outstation -				
the European Bioinformatics Institute. There are no restrictions on its				
use by non-profit institutions as long as its content is in no way				
modified and this statement is not removed. Usage by and for commercial				
entities requires a license agreement (See http://www.isb-sib.ch/announce/				
or send an email to license@isb-sib.ch).				

EMBL; M84488; AAA42332.1; -.				
PIR; J50675; J50675.				
HSSP; P19320; IVCA.				
InterPro; IPR003987; ICAM VCAM-1.				
InterPro; IPR007110; Ig-like.				
InterPro; IPR003598; Ig C2.				
InterPro; IPR003989; VCAM-1.				
Pfam; PF00047; Ig; 5.				
PRINTS; PR01472; ICAMVCAM1.				
PRINTS; PR01474; VCAM1.				
SMART; SM00408; IGC2; 3.				
PROSITE; PS50835; IG_LIKE; 5.				
Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;				
Repeat; Signal.				
FT SIGNAL 1 24 PROBABLE.				
FT CHAIN 25 739 VASCULAR CELL ADHESION PROTEIN 1.				
FT DOMAIN 25 698 EXTRACELLULAR (POTENTIAL).				
FT TRANSMEM 699 720 POTENTIAL.				
FT DOMAIN 721 739 CYTOPLASMIC (POTENTIAL).				
FT DOMAIN 25 111 IG-LIKE C2-TYPE 1.				
FT DOMAIN 119 212 IG-LIKE C2-TYPE 2.				
FT DOMAIN 223 309 IG-LIKE C2-TYPE 3.				
FT DOMAIN 312 397 IG-LIKE C2-TYPE 4.				
FT DOMAIN 408 506 IG-LIKE C2-TYPE 5.				
FT DOMAIN 514 595 IG-LIKE C2-TYPE 6.				
FT DOMAIN 601 682 IG-LIKE C2-TYPE 7.				
FT DISULFID 47 95 BY SIMILARITY.				
FT DISULFID 52 99 BY SIMILARITY.				
FT DISULFID 137 195 BY SIMILARITY.				
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT CARBOHYD 424 424 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).				
FT CARBOHYD 650 650 N-LINKED (GLCNAC. . .) (POTENTIAL).				
SEQUENCE 739 AA; 81246 MW; 5608E5A1A1B100C CRC64;				

Best Local Similarity 21.4%; Pred. No. 0.00059;
Matches 101; Conservative 71; Mismatches 169; Indels 131; Gaps 23;

```

QY 33 KGDYELTQTASQKSIQPHMKNQNDIKLNGSFLTGPBKLANDRADRSRLMDQNF 92
DB 238 EAAAVTMTCTASGSLPAPPELFMSKK-----LDNGVQLL-----SSNA 274
QY 93 PLIKNKLIKEDSDTYICE---VEDQKEVQLLV-----FGLTANSPDTHLLOGLTLT 142
DB 275 TLTLLAMREBDSGIVYCEGVNLVGDKTEVEILVQEKPTVDISGSGVAAQVGSVILT 334
QY 143 LESPGSSPSVOCSPRGKNIGQ-----GKTLVSQLELDSDGTWTCTVLQNGKYVF 195
DB 335 CAACVCDSPSFGMRQTQDPLNGEVRDEGATSTLTLSIPGVDEHSYLTCTVQGRKLEK 394
QY 196 KLDIYPCAPAPKSCDKTKHTCEPL-LGSPSVLFPKPKDPTLMISTREYVCVVYD--- 251
DB 395 TIQVEYVSFPD-----PEIETSGP-----LVHGRPTVNTCTVBNVYPF 433
QY 252 SHEDEEV-----KFNMYVDGVEVHNAKTKPREEQY---NSTYRVSVLTVLHQDW 298
DB 434 DLLELELKGFTTLNKF--LAEIGTSLFETKSLDMPTFPAEDTGALVCLATHSSQ 491
QY 299 LNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLR-PSRDELTKNQVSLTCLVKGF 357
DB 492 ME-SEPKQKQSTQTLVYVNV-----APKEPTIWSPPSPVPER--GSPVNLTCSSDGF 539
QY 358 YPSDAVEME---SNGQPENNYKTPPVLDSDGSFVLSKLTVDKSRNQGVFSCSYMH 414
DB 540 PTPKLL--WSROLKNGELQ-----PLSQNTTSLFPMATKM-----EDSGIYVCEGIN 583
QY 415 EALAHNYTQKSLSL-----SPGLQL-----DETCAEAGDGLGLM 450
DB 584 EA---GISKKSVELLIIGSSKDIQLTVFSPSKSVKGDVYIISCTGSGVPEIM 632

```

```

RESULT 65
ID PGEM HUMAN STANDARD; PRT; 4391 AA.
AC P98160; O16287; Q9H3V5;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DB 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (Perlecan) (PLC).
GN HSPG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=92112994; PubMed=1730768;
RA Kallunki P., Tryggvason K.;
RT "Human basement membrane heparan sulfate proteoglycan core protein: a
RT 467-kD protein containing multiple domains resembling elements of the
RT low density lipoprotein receptor, laminin, neural cell adhesion
RT molecules, and epidermal growth factor."
RL J. Biol. Chem. 267:8544-8557(1992).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=92235084; PubMed=1569102;
RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;
RT "Primary structure of the human heparan sulfate proteoglycan from
RT basement membrane (HSPG2/perlecan). A chimeric molecule with multiple
RT domains homologous to the low density lipoprotein receptor, laminin,
RT neural cell adhesion molecules, and epidermal growth factor."
RL J. Biol. Chem. 267:8544-8557(1992).
RN 13
RP SEQUENCE OF 22-4391 FROM N.A., AND VARIANT SJS1 TYR-1532.
RX MEDLINE=20553141; PubMed=11101850;
RA Nicole S., Davoine C.-S., Topaloglu H., Catolico L., Barral D.,

```

```

RA Beighton P., Ben-Hamida C., Hammouda H., Craud C., White P.S.,
RA Sanson D., Urtizberea J.A., Lehmann-Horn F., Weissenbach J.,
RA Hentati F., Fontaine B., Lehmann-Horn F., Weissenbach J.,
RT "Perlecan, the major proteoglycan of basement membranes, is altered in
RT patients with Schwartz-Jampel syndrome (chondrodysdystrophic myotonia).";
RL Nat. Genet. 26:480-483(2000).
RN 14
RP SEQUENCE OF 1016-1470 FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=91365376; PubMed=1679749;
RA Dodge G.R., Kovalszky I., Chu M.L., Hasell J.R., McBride O.W.,
RA Yi H.F., Iozzo R.V.;
RT "Heparan sulfate proteoglycan of human colon: partial molecular
RT cloning, cellular expression, and mapping of the gene (HSPG2) to the
RT short arm of human chromosome 1."
RL Genomics 10:673-680(1991).
RN 15
RP SEQUENCE OF 890-1396 FROM N.A.
RC TISSUE=Fibrocarcoma;
RX MEDLINE=92120660; PubMed=1685141;
RA Kallunki P., Eddy R.L., Byers M.G., Keestila M., Shows T.B.,
RA Tryggvason K.;
RT "Cloning of human heparan sulfate proteoglycan core protein,
RT assignment of the gene (HSPG2) to 1p36.1-1p35 and identification of
RT a BamHI restriction fragment length polymorphism."
RL Genomics 11:389-396(1991).
RN 16
RP SEQUENCE OF 1-21 FROM N.A.
RX MEDLINE=94052171; PubMed=8234307;
RA Cohen I.R., Graessell S., Murdoch A.D., Iozzo R.V.;
RT "Structural characterization of the complete human perlecan gene and
RT its promoter."
RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408(1993).
RN 17
RP CARBOHYDRATE-LINKAGE SITE ASN-2121.
RX MEDLINE=22660472; PubMed=12754519;
RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT "Identification and quantification of N-linked glycoproteins using
RT hydrazide chemistry, stable isotope labeling and mass spectrometry."
RL Nat. Biotechnol. 21:660-666(2003).
RN 18
RP FUNCTION: This protein is an integral component of basement
RN membranes. It is responsible for the fixed negative electrostatic
RN charge and is involved in the charge-selective ultrafiltration
RN properties. It serves as an attachment substrate for cells.
RN 19
RP SUBUNIT: Purified perlecan has a strong tendency to aggregate in
RN dimers or stellate structures. It interacts with other basement
RN membrane components such as laminin, prolargin and collagen type
RN IV.
RN 20
RP SUBCELLULAR LOCATION: Extracellular.
RN 21
RP TISSUE SPECIFICITY: Found in the basement membranes.
RN 22
RP -1- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
RN AND O-LINKED OLIGOSACCHARIDES.
RN 23
RP -1- DISEASE: Defects in HSPG2 are the cause of Schwartz-Jampel
RN syndrome (SJS1) [MIM:255800]; a rare autosomal recessive disorder
RN characterized by permanent myotonia (prolonged failure of muscle
RN relaxation) and skeletal dysplasia, resulting in reduced stature,
RN kyphoscoliosis, bowing of the diaphyses and irregular epiphyses.
RN 24
RP -1- SIMILARITY: Contains 4 LDL-receptor class A domains.
RN 25
RP -1- SIMILARITY: Contains 11 laminin EGF-like domains.
RN 26
RP -1- SIMILARITY: Contains 3 laminin IV domains.
RN 27
RP -1- SIMILARITY: Contains 22 immunoglobulin-like C2-type domains.
RN 28
RP -1- SIMILARITY: Contains 4 EGF-like domains.
RN 29
RP -1- SIMILARITY: Contains 1 SEA domain.
RN 30
RP -----
RN 31
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.ebi.ac.uk/announcements
RN or send an email to license@ebi.ac.uk).
RN -----

```

```

DR EMBL: X62515; CAA44373.1; -.
DR EMBL: M85289; AAA52700.1; -.
DR EMBL: AL445795; CAC18534.1; -.
DR EMBL: M64283; AAA52699.1; -.
DR EMBL: S76436; AAB2121.2; -.
DR EMBL: L22078; -, NOT_ANNOTATED_CDS.
DR PIR: A38096; A38096.
DR HSSP: P00740; 1EDM.
DR Siema-2DPAGE; P98160; -.
DR Genew; HGNC:3273; HSP62.
DR MIM; 142461; -.
DR MIM; 255800; -.
DR InterPro: IPR008985; ConA_like_1ec_g1.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003596; Ig_v.
DR InterPro: IPR000034; Laminin_B.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR001791; Laminin_G.
DR InterPro: IPR002172; LDL_receptor_A.
DR Pfam; PPR00082; SEA_domain.
DR Pfam; PR00047; EGF_4.
DR Pfam; PR00052; Laminin_B_3.
DR Pfam; PR00053; Laminin_EGF_7.
DR Pfam; PR00054; Laminin_G_3.
DR Pfam; PR00057; ldl_recept_a; 4.
DR Pfam; PR01390; SEA_1.
DR PRINTS; PR00261; LDLRECEPTOR.
DR ProDom; PD003031; Laminin_B_3.
DR SMART; SM00181; EGF_15.
DR SMART; SM00180; EGF_Lam; 12.
DR SMART; SM00409; IG_22.
DR SMART; SM00409; IGC2; 21.
DR SMART; SM00406; IGV; 7.
DR SMART; SM00281; Lamb; 3.
DR SMART; SM00282; Lamg; 3.
DR SMART; SM00192; LDLA; 4.
DR SMART; SM00200; SEA; 1.
DR PROSITE; PS00022; EGF_1; 9.
DR PROSITE; PS01186; EGF_2; 6.
DR PROSITE; PS50026; EGF_3; 4.
DR PROSITE; PS50835; IG_Like; 22.
DR PROSITE; PS50025; LAM_G_DOMAIN; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
DR PROSITE; PS01209; LDLRA_1; 4.
DR PROSITE; PS50068; LDLRA_2; 4.
DR PROSITE; PS50024; SEA; 1.
KW Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;
KW Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;
KW Extracellular matrix; EGF-like domain; Disease mutation.
FT SIGNAL 1 21
FT CHAIN 22 4391
FT DOMAIN 80 194
FT DOMAIN 198 235
FT DOMAIN 284 320
FT DOMAIN 324 360
FT DOMAIN 367 404
FT DOMAIN 405 504
FT DOMAIN 521 530
FT DOMAIN 531 730
FT DOMAIN 731 763
FT DOMAIN 764 813
FT DOMAIN 814 871
FT DOMAIN 879 933
FT DOMAIN 924 933
FT DOMAIN 934 1125
FT DOMAIN 1126 1158

```

```

FT DOMAIN 1159 1208 LAMININ EGF-LIKE 6.
FT DOMAIN 1209 1265 LAMININ EGF-LIKE 7.
FT DOMAIN 1275 1324 LAMININ EGF-LIKE 8.
FT DOMAIN 1325 1334 LAMININ EGF-LIKE 9 (N-TERMINAL).
FT DOMAIN 1335 1529 LAMININ DOMAIN IV 3 (DOMAIN III C).
FT DOMAIN 1530 1562 LAMININ EGF-LIKE 9 (C-TERMINAL).
FT DOMAIN 1563 1612 LAMININ EGF-LIKE 10.
FT DOMAIN 1613 1670 LAMININ EGF-LIKE 11.
FT DOMAIN 1671 1771 IG-LIKE C2-TYPE 2.
FT DOMAIN 1772 1865 IG-LIKE C2-TYPE 3.
FT DOMAIN 1866 1955 IG-LIKE C2-TYPE 4.
FT DOMAIN 1956 2051 IG-LIKE C2-TYPE 5.
FT DOMAIN 2052 2151 IG-LIKE C2-TYPE 6.
FT DOMAIN 2152 2244 IG-LIKE C2-TYPE 7.
FT DOMAIN 2245 2340 IG-LIKE C2-TYPE 8.
FT DOMAIN 2341 2436 IG-LIKE C2-TYPE 9.
FT DOMAIN 2437 2533 IG-LIKE C2-TYPE 10.
FT DOMAIN 2534 2629 IG-LIKE C2-TYPE 11.
FT DOMAIN 2630 2726 IG-LIKE C2-TYPE 12.

Query Match 6.5% Score 175.5, DB 1; Length 4391;
Best Local Similarity 20.7%; Pred. No. 0.0064;
Matches 100; Conservative 67; Mismatches 194; Indels 123; Gaps 22;

QY 33 KSDTVELCTASQKSIQPHMKNNSQIKILNGSFLKRGPSKLNDRADSRSLMDQGNF 92
DB 2740 ESETLDLNCVVPQGAHQVTH-----KRGSL--PSHHOTRGRRL----- 2779
QY 93 PLIRLNKIEDSDTYICEVEDK--EEVOLVFGILTANSDTH----- 132
DB 2780 ---LHHVSPADSGEYVCVWSSGSLASVLTITASGSAVHVPAGAPRIEPPSS 2836
QY 133 -LLQGSLTLTIESPPSSPSVQCSPRGKNIQG-----GKTLVSQLELDGSGTWTCV 186
DB 2837 RVAEQTLDLKC-VVPGQAHQVTHKRGKGNLPAHQVHGPLRLNQVSPADSGEYSCV 2895
QY 187 LONQKVEPKIDIV-----PCPAPPKSGCDKHTCPRELLGSPVLPFPKPKDTLMISRT 241
DB 2896 TGSSGTLASVLTIEPPSPGPIPAV-----GLAQPIYBASSHYTREGT 2941
QY 242 PEVTCVVVDVSHEDPEVFKNFQVDEVEYHNAKTKRREQYNSTYVSVLTVLHODWING 301
DB 2942 LDLCNVVPGQH---AQVTWYKRG-----GSLPRHQTHGQLRLHLVSPA-----DS 2986
QY 302 KEYKCVSNKALP--APIETKISAKQOP--REPQVYTLPPSDELTAKNOVSITCLV- 354
DB 2987 GEYVCRAASGPQEQEASFVTVPPSEGSYRLRPVISIDPPSSSTVOQGDASFKCLIH 3046
QY 355 KGFYPSDLAVENESGQP-ENNYKTPPVLDSDGSFFLYSKULTVDSKMQGQNVFSC--- 410
DB 3047 DGAAP--ISLEWKTNNQLELDVHISP-----NGSI-----ITVGTBPSNHTYRCVAS 3094
QY 411 -----SYMHEALNHYTKSLISLP-----GLQLDTCFAQDGEIDGLTTDPPRAS 458
DB 3095 NAYGVAQGVNLSVIGPPTVSVLPBGVPVKKVKAVTIECVSAGSPRSARAT-----RIS 3150
QY 459 ALPA 462
DB 3151 STPA 3154

RESULT 66
LAC2_RAT STANDARD; PRT; 104 AA.
AC P20767;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig lambda-2 chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87055594; PubMed=3114047;
RA Steen M.L., Hellman L., Peterson U.;
RT "The immunoglobulin lambda locus in rat consists of two C lambda
RT genes and a single V lambda gene."
RL CC
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
CC EMBL, M22521; AAA1420.1; ALT_INIT.
DR HSP; P01842; 2MCG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG_1.
DR SMART; SM00407; IGc1.1.
DR PROSITE; PS50835; IG-LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR Immunoglobulin domain; Immunoglobulin C region.
FT NON TER 1 1
FT DOMAIN 6 99
FT DISULFID 27 85
FT DISULFID 103 103
SQ SEQUENCE 104 AA; 11318 MW; F087906DE43F7276 CRC64;

Query Match
Best Local Similarity 35.3%; Score 172; DB 1; Length 104;
Matches 36; Conservative 17; Mismatches 45; Indels 4; Gaps 2;

QY 327 QPRE-POVYTLPPSRDELTKQVSLTCLVKGFPSDIAVEMESNQPPENNYKTPPVVDS 385
DB 1 QPKSPFTLVPPSPSTBELQGNKATLVCLISDFPSDEVAMKANCAPISQGVDTIANPKQ 60

QY 386 DGSFPLYSKLTVDKSRMOQGNVSCSVMEALHNHYTKSKIS 427
DB 61 GNKYIASFFLRRLTAQWRSRNSFTQVTHE--GNTVERKSLIS 99

RESULT 67
DTC_HUMAN STANDARD; PRT; 383 AA.
ID DTC_HUMAN
AC P01880;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig delta chain C region.
GN IGHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBT_Taxid=9606;
RX MEDLINE=82082419; PubMed=6947220;
RX Putnam F.W., Takahashi N., Tetaert D., Debuire B., Lin L.-C.;
RT "Amino acid sequence of the first constant region domain and the
RT hinge region of the delta heavy chain of human IgD."
RL Proc. Natl. Acad. Sci. U.S.A. 78:6168-6172(1981).
RN [2]
RP SEQUENCE OF 158-383 (MYELOMA PROTEIN MAH).
RX MEDLINE=81223768; PubMed=6787589;
RA Lin L.-C., Putnam F.W.;
RT "Primary structure of the C region of human immunoglobulin D:
RT implications for evolutionary origin and biological function."
RL Proc. Natl. Acad. Sci. U.S.A. 78:504-508(1981).

```

```

RN [3]
RP SEQUENCE OF 103-137 (MYELOMA PROTEIN NIG-65), AND CARBOHYDRATES.
RX MEDLINE=82231401; PubMed=7092891;
RA Takayasu T., Suzuki S., Kametani F., Takahashi N., Shinoda T.,
RA Okuyama T., Munekeata E.;
RT "Amino acid sequence of galactosamine-containing glycopeptides in the
RT hinge region of a human immunoglobulin D."
RL Biochem. Biophys. Res. Commun. 105:1066-1071(1982).
RN [4]
RP SEQUENCE OF 158-383 (MYELOMA PROTEIN NIG-65).
RX MEDLINE=81199406; PubMed=6785754;
RA Shinoda T., Takahashi N., Takayasu T., Okuyama T., Shimizu A.;
RT "Complete amino acid sequence of the C region of a human
RT chain."
RL Proc. Natl. Acad. Sci. U.S.A. 78:785-789(1981).
RN [5]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=81133632; PubMed=7008791;
RA Takayasu T., Takahashi N., Shinoda T.;
RT "Amino acid sequence and location of the three glycopeptides in the
RT C region of human immunoglobulin D."
RL Biochem. Biophys. Res. Commun. 97:635-641(1980).
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
DR PIR; A02175; DHU.
DR Genew; HGNC:5480; IGHD.
DR MIM; 147170; -.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG_3.
DR SMART; SM00407; IGc1.3.
DR PROSITE; PS50835; IG-LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 2.
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
FT NON TER 1 1
FT DOMAIN 6 98
FT DISULFID 175 263
FT DISULFID 267 373
FT DISULFID 15 15
FT DISULFID 28 84
FT DISULFID 161 161
FT CARBOHYD 109 109
FT CARBOHYD 110 110
FT CARBOHYD 113 113
FT CARBOHYD 126 126
FT CARBOHYD 127 127
FT CARBOHYD 131 131
FT CARBOHYD 132 132
FT CARBOHYD 225 225
FT CARBOHYD 316 316
FT CARBOHYD 367 367
SQ SEQUENCE 383 AA; 42125 MW; DA1DE73519C76CID CRC64;

Query Match
Best Local Similarity 23.1%; Score 170; DB 1; Length 383;
Matches 83; Conservative 41; Mismatches 118; Indels 118; Gaps 16;

QY 232 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYGVGFEVHNAAKTRPEEQYNSTYRVVSVL 291
DB 18 PKD-----NSPVYLACLI-TGYHPTSVTVWYWM-GTQSQGQRFPEIQRDSYVMTSSQL 70

QY 292 TVLHQMVLNKEVKKYKSN----- 316
DB 71 STPLQMRQG-EKCVQVHTASKSKKEIFRWPSPKQAQSSVPTPAQAGSLAKATTPAP 129

QY 317 I-----EKTISKAKGQPREPQ-----VYTLPPSRDEL-TKNQVSLT 351
DB 130 ATRNRNGRGEKKKEKKKEKEKEKEKEKEKEKEKESHQPLQGLVLTTPAVQDLMLRDKATFT 189

QY 352 CLVKGFPSPDIAVEMESNQ-PENNYKTPPVVDSGSPFLYSKLTVDKSRMOQGNVSC 410

```


DB 190 CFVWGSDLDKDALTWLWVACKVPTGVEEGLLEHRSNGSQSHSRLLTPSLMNAAGTSTVC 249
QY 411 SYWHEALHNYTKSLSPG--LQUDFCAEAG-DGELDGLMTTPPASAALP----- 461
DB 250 TLNHP-----SLPPQRLMALREPAQAQPVKLSLNTLASSDPPPEAASWLCESVG 298
QY 462 -APPT-----GSALPDPQTASA-----LPDPPEASALPALAVI 494
DB 299 FSPBNLMLWLEQREVNTSGFAPRPPQPGSTTWMASVLRVPAP--SQPAPYTTCV 356
RESULT 68
CD22_HUMAN STANDARD: PRT: 847 AA.
ID P20273; Q95699; Q95701; Q95702; Q95703; Q01665; Q92872; Q92873;
AC Q9UQA7; Q9UQA8; Q9UQA9; Q9UQB0; Q912A6;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE B-cell receptor CD22 precursor (Leu-14) (B-lymphocyte cell adhesion molecule) (BL-CAM) (Siglec-2).
GN CD22.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1] SEQUENCE FROM N.A. (ISOFORM CD22-BETA), AND VARIANT HIS-639.
RC TISSUE=Toneil;
RX MEDLINE=91086838; PubMed=1985119;
RA Wilson G.L., Fox C.H., Fauci A.S., Kehrl J.H.;
RT "CDNA cloning of the B cell membrane protein CD22: a mediator of B-B cell interactions.";
J. Exp. Med. 173:137-146(1991).
RN [2] SEQUENCE FROM N.A. (ISOFORMS CD22-ALPHA AND CD22-BETA).
RX MEDLINE=93267103; PubMed=8496602;
RA Wilson G.L., Neifeld V., Kozlow E., Menninger J., Ward D.,
RT Kehrl J.H.;
RN "Genomic structure and chromosomal mapping of the human CD22 gene.";
J. Immunol. 150:5013-5024(1993).
RN [3] SEQUENCE FROM N.A. (ISOFORMS CD22-ALPHA AND CD22-BETA).
RX Lamerding J.E., McCreedy P., Adamson A.W., Burkhardt-Schultz K.,
RA Garcia E., Kyle A., Ramirez M., Stillwagen S., Garnea J.,
RA Dangnanan L., Bruce R., Quan G., Montgomery W., Ow D.,
RA Kobayashi A., Olsen A.O., Carraro A.V.;
RN Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
RN [4] SEQUENCE FROM N.A. (ISOFORM CD22-BETA).
RX MEDLINE=90231465; PubMed=1691828;
RA Stamenkovic I., Seed B.;
RN "The B-cell antigen CD22 mediates monocyte and erythrocyte adhesion.";
Nature 345:74-77(1990).
RN [5] SEQUENCE OF 13-137, 139-239, 241-328 AND 418-502 FROM N.A., AND
RX VARIANTS THR-34; GLU-152; LYS-203; GLY-664; CYS-669 AND ASP-745.
RN MEDLINE=93180618; PubMed=10079291;
RA Hata Y., Tsuchiya N., Matsushita M., Shiota M., Hagiwara K.,
RA Tokunaga K.;
RN "Identification of the gene variations in human CD22.";
Immunogenetics 49:280-286(1999).
RN [6] SIALIC ACID BINDING.
RX MEDLINE=93216636; PubMed=8463235;
RA Powell L.D., Segol D., Sjoberg E.R., Stamenkovic I., Varki A.;
RN "Natural ligands of the B cell adhesion molecule CD22 beta carry N-linked oligosaccharides with alpha-2,6-linked sialic acids that are required for recognition.";
J. Biol. Chem. 268:7019-7027(1993).
RN [7] INTERACTION WITH PTPN6.
RX MEDLINE=95343349; PubMed=7618087;

RA Doody G.M., Justement L.B., Delibrias C.C., Matthews R.J., Lin J.,
RA Thomas M.L., Featon D.T.;
RN "A role in B cell activation for CD22 and the protein tyrosine phosphatase SHP.";
Science 269:242-244(1995).
RN [8] INTERACTION WITH LYN, SYK AND PIK3R1/PIK3R2.
RX MEDLINE=96257803; PubMed=8647200;
RA Tuscano J.M., Engel P., Tedder T.F., Agarwal A., Kehrl J.H.;
RN "Involvement of p72syk kinase, p53/56lyn kinase and phosphatidylinositol-3 kinase in signal transduction via the human B lymphocyte antigen CD22.";
Eur. J. Immunol. 26:1246-1252(1996).
RN [9] INTERACTION WITH PTPN6, SYK AND PLCG1.
RX MEDLINE=96195207; PubMed=8627166;
RA Law C.L., Sidorenko S.P., Chandran K.A., Zhao Z., Shen S.H.,
RA Fischer B.H., Clark E.A.;
RN "CD22 associates with protein tyrosine phosphatase 1C, Syk, and phospholipase C-gamma(1) upon B cell activation.";
J. Exp. Med. 183:547-560(1996).
RN [10] REVIEW.
RX MEDLINE=97288746; PubMed=9143697;
RA Tedder T.F., Tuscano J., Sato S., Kehrl J.H.;
RN "CD22, a B lymphocyte-specific adhesion molecule that regulates antigen receptor signaling.";
Annu. Rev. Immunol. 15:481-504(1997).
RN [11] FUNCTION: Mediates B-cell B-cell interactions. May be involved in the localization of B-cells in lymphoid tissues. Binds sialylated alpha2,6-linked sialic acid. The sialic acid recognition site can be masked by cis interactions with sialic acids on the same cell surface. Upon ligand induced tyrosine phosphorylation in the immune response seems to be involved in regulation of B cell antigen receptor signaling. Plays a role in positive regulation through interaction with Src family tyrosine kinases and may also act as an inhibitory receptor by recruiting cytoplasmic phosphatases via their SH2 domains that block signal transduction through dephosphorylation of signaling molecules.
CC [1] SUBUNIT: Predominantly monomer of isoform CD22-beta. Also found as heterodimer of isoform CD22-beta and a shorter isoform. Interacts with PTPN6/SHP-1, LYN, SYK, PIK3R1/PIK3R2 and PLCG1 upon phosphorylation. Interacts with GRB2 and SHC1 upon phosphorylation (By similarity). May form a complex with INP5D/SHP, GRB2 and SHC1.
CC [1] SUBCELLULAR LOCATION: Type I membrane protein.
CC [1] ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=CD22-beta;
CC IsoId=P20273-1; Sequence=Displayed;
CC Name=CD22-alpha;
CC IsoId=P20273-2; Sequence=VSP_002531;
CC [1] TISSUE SPECIFICITY: B lymphocytes.
CC [1] DOMAIN: Contains 4 copies of a cytoplasmic motif that is referred to as the immunoreceptor tyrosine-based inhibitor motif (ITIM). This motif is involved in downmodulation of cellular responses. The phosphorylated ITIM motif binds to the SH2 domain of PTPN6/SHP-1.
CC [1] PTM: Phosphorylated both on threonine/serine and tyrosine residues.
CC [1] PTM: Phosphorylated on tyrosine residues by LYN (By similarity).
CC [1] SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC (sialic acid binding Ig-like lectin) family.
CC [1] SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC [1] SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC [1] CAUTION: Ref.4 sequence differs from that shown due to a frameshift in position 806.
CC [1] DATABASE: NAME=PROW; NOTE=CD guide CD22 entry; htm.
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd22.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.B.,
RA Swann R.M., Symamoe N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Skin;
RA MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heist F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uebin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield J.S.N., Krzywinski M.I., Skalek U., Smalleg D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP INTERACTION WITH TYROBP AND SYK.
RA PubMed=11169422;
RA Tomaseello E., Cant C., Buehring H.-J., Vely F., Andre P., Seifert M.,
RA Ullrich A., Vavler E.;
RT "Association of signal-regulatory protein beta with KARAP/DAP-12.";
RL Eur. J. Immunol. 30:2147-2156(2000).
RN [5]
RP INTERACTION WITH TYROBP, N-GLYCOSYLATION, AND TISSUE SPECIFICITY.
RX MEDLINE=20072721; PubMed=10604985;
RA Dietrich J., Cella M., Seifert M., Buehring H.-J., Colonna M.;
RT "Signal-regulatory protein beta 1 is a DAP12-associated activating
RT receptor expressed in myeloid cells.";
RL J. Immunol. 164:9-12(2000).
CC -1- FUNCTION: Immunoglobulin-like cell surface receptor involved in
CC the negative regulation of receptor tyrosine kinase-coupled
CC signaling processes. Participates also in the recruitment of
CC tyrosine kinase SYK.
CC -1- SUBUNIT: Interacts with TYROBP. This interaction results in the
CC recruitment of SYK.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O00241-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O00241-2; Sequence=VSP_007026;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Detected in monocytes and dendritic cells.
CC -1- PTM: N-glycosylated.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y10376; CAA71404.1; -
CC EMBL: AL049634; CAB46661.2; -
CC EMBL: AL13804; CAC17540.1; -
CC EMBL: BC025286; AAH25286.1; -
CC Genew: HGNC:15928; STRPB1.
CC MIM: 603889;
DR GO:0005887; C: integral to plasma membrane; TAS.
DR GO:0007166; P: cell surface receptor linked signal transdu. .; TAS.
DR InterPro: IPR007110; IG-1-like.
DR InterPro: IPR003597; IG_c1.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00407; Ig_c1_2.
DR PROSITE: PS00835; IG_LIKE_3.
KW Repeat; Signal; Transmembrane; Immunoglobulin domain; Glycoprotein;
KW Alternative splicing.
FT SIGNAL 1 26
FT CHAIN 27 398
FT DOMAIN 27 371
FT TRANSMEM 372 392
FT DOMAIN 393 398
FT DOMAIN 27 136
FT DOMAIN 147 246
FT DOMAIN 253 347
FT DISULFID 34 120
FT DISULFID 169 227
FT DISULFID 272 330
FT CARBOHYD 244 244
FT CARBOHYD 244 244
FT CARBOHYD 269 269
FT CARBOHYD 291 291
FT VARSPLIC 145 361
FT CONFLICT 53 53
FT CONFLICT 102 102
FT CONFLICT 229 229
FT CONFLICT 363 363
SQ SEQUENCE 398 AA; 43255 MW; A2AA08FE8B2BC52B CRC64;
Query Match 6.2%; Score 166.5; DB 1; Length 398;
Best Local Similarity 24.2%; Pred. No. 0.0012;
Matches 82; Conservative 50; Mismatches 154; Indels 53; Gaps 14;
QY 136 GGSLLTTPSPGSSDPSVQCRKNTQGGKTLVSQLELDSGTWTGTLQ--NQKV 193
DB 72 GAGRELIYNQKKGHPRTVTTSSELTRNNLDFISISINITPADAGTYCVCKFRKGSDDV 131
QY 194 EKKIDVPCPAPBPSCDKTHCPRLGSPSYFLPPPKKDLMTSRPEVYCVVVDVSH 253
DB 132 EFK----SGAGTSLVRAKPPAP--VSGPAP-----RAPEHTVSTCESH 172
QY 254 --EDPEKFNMYVDVGVENAKTK--PREE-----QYNSTYRVSVTLVLDHDLNGKEYK 305
DB 173 GSPSPDITLKWFKNGKSLSDPQTNNVDPAGDSYSHSTARVYLRGDVHSGVI----- 226
QY 306 CYSNKAFLPA-DIEKT--ISKAGQPREQVYTLPPSRDELTKQVSLTCLVKGFPYSDI 362
DB 227 CEMAHITLQGDPLRGANLSEALRVPLEV--TQCPMRAE---NOANVTQGVSNFPRGL 282
QY 363 AVESWNSQNPENNYKTPPVLDSDGSFLYSLTLYDQSKWQGNVPSGCVMEALHNHT 422
DB 283 QLTWLENGVNSRTETASTLIENKDKGYMMGSLVNTCAHRDQVVLTCVEHDG----- 336
QY 423 QKSLSPGLQDETCAEKNODELGLMTTDPPRASALP 461
DB 337 QQAVSKSYALEI-----SAHQEHGSDITHEPALPTAP 370
RESULT 70
UN89 CAEEL
ID UN89 CAEEL STANDARD; PRT; 6632 AA.
AC 001761; Q17362;

DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).
 GN UNC-89 OR C09D1.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC STRAIN-Bristol N2;
 RC MEDLINE=96180278; PubMed=8603916;
 RA Benjan G.M., Tinley T.L., Tang X., Borodovsky M.;
 RT "The Caenorhabditis elegans gene unc-89, required for muscle M-line
 RT assembly, encodes a giant modular protein composed of Ig and signal
 RT transduction domains.";
 RT J. Cell Biol. 132:835-848 (1996).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-Bristol N2;
 RC Du Z., Le T.T., Wilson R.;
 RL Submitted (May-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Structural component of the muscle M-line. Myofibril
 CC lattice assembly begins with positional cues laid down in the
 CC basement membrane and muscle cell membrane. UNC-89 responds to
 CC these signals, localizes, and then participates in assembling an
 CC M-line.
 CC -1- TISSUE SPECIFICITY: Localizes to the middle of A-bands.
 CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type III domain.
 CC -1- SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -1- SIMILARITY: Contains 5 RCPD domains.
 CC -1- SIMILARITY: Contains 1 SH3 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U33058; AAB00542.1; -.
 DR EMBL: AF003131; AAB54132.2; -.
 DR PDB: 1PHO; 20-DEC-00.
 DR WormPep: C09D1.1; CE30426.
 DR InterPro: IPR008957; FN_III-like.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR007110; IG-1-like.
 DR InterPro: IPR003598; IG_C2.
 DR InterPro: IPR003006; IG_MHC.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR007850; RCPD.
 DR InterPro: IPR000219; RNOGEF.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF000441; fn3; 1.
 DR Pfam: PF000447; Ig; 47.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00177; RCPD; 5.
 DR Pfam: PF00621; RNOGEF; 1.
 DR Pfam: PF00018; SH3; 1.
 DR SMART: SMO0408; IG_C2; 23.
 DR SMART: SMO0325; RNOGEF; 1.
 DR SMART: SMO0326; SH3; 1.
 DR PROSITE: PS50010; DH_2; 1.
 DR PROSITE: PS50835; IG_LIKE; 49.
 DR PROSITE: PS50003; PH_DOMAIN; 1.

DR PROSITE: PS50002; SH3; 1.
 KW Muscle protein; Immunoglobulin domain; Repeat; SH3 domain;
 KW 3D-structure.
 FT DOMAIN 63 127 SH3.
 FT DOMAIN 152 330 DH.
 FT DOMAIN 342 498 PH.
 FT DOMAIN 547 633 IG-LIKE C2-TYPE 1.
 FT DOMAIN 648 736 IG-LIKE C2-TYPE 2.
 FT DOMAIN 748 838 IG-LIKE C2-TYPE 3.
 FT DOMAIN 946 1033 IG-LIKE C2-TYPE 4.
 FT DOMAIN 1044 1132 IG-LIKE C2-TYPE 5.
 FT DOMAIN 1140 1227 IG-LIKE C2-TYPE 6.
 FT DOMAIN 1272 1315 THR-RICH.
 FT DOMAIN 1375 1475 RCPD 1.
 FT DOMAIN 1479 1585 RCPD 2.
 FT DOMAIN 1597 1695 RCPD 3.
 FT DOMAIN 1700 1799 RCPD 4.
 FT DOMAIN 1800 1860 RCPD 5.
 FT DOMAIN 1982 2067 IG-LIKE C2-TYPE 7.
 FT DOMAIN 2071 2163 IG-LIKE C2-TYPE 8.
 FT DOMAIN 2171 2261 IG-LIKE C2-TYPE 9.
 FT DOMAIN 2269 2359 IG-LIKE C2-TYPE 10.
 FT DOMAIN 2367 2455 IG-LIKE C2-TYPE 11.
 FT DOMAIN 2463 2564 IG-LIKE C2-TYPE 12.
 FT DOMAIN 2563 2651 IG-LIKE C2-TYPE 13.
 FT DOMAIN 2657 2746 IG-LIKE C2-TYPE 14.
 FT DOMAIN 2754 2858 IG-LIKE C2-TYPE 15.
 FT DOMAIN 2887 2980 IG-LIKE C2-TYPE 16.
 FT DOMAIN 2894 3081 IG-LIKE C2-TYPE 17.
 FT DOMAIN 3087 3183 IG-LIKE C2-TYPE 18.
 FT DOMAIN 3189 3280 IG-LIKE C2-TYPE 19.
 FT DOMAIN 3286 3376 IG-LIKE C2-TYPE 20.
 FT DOMAIN 3384 3469 IG-LIKE C2-TYPE 21.
 FT DOMAIN 3482 3572 IG-LIKE C2-TYPE 22.
 FT DOMAIN 3580 3667 IG-LIKE C2-TYPE 23.
 FT DOMAIN 3686 3777 IG-LIKE C2-TYPE 24.
 FT DOMAIN 3817 3908 IG-LIKE C2-TYPE 25.
 FT DOMAIN 3920 4009 IG-LIKE C2-TYPE 26.
 FT DOMAIN 4018 4106 IG-LIKE C2-TYPE 27.
 FT DOMAIN 4109 4201 IG-LIKE C2-TYPE 28.
 FT DOMAIN 4212 4297 IG-LIKE C2-TYPE 29.
 FT DOMAIN 4302 4387 IG-LIKE C2-TYPE 30.
 FT DOMAIN 4400 4485 IG-LIKE C2-TYPE 31.
 FT DOMAIN 4489 4580 IG-LIKE C2-TYPE 32.
 FT DOMAIN 4588 4678 IG-LIKE C2-TYPE 33.
 FT DOMAIN 4681 4771 IG-LIKE C2-TYPE 34.
 FT DOMAIN 4873 4961 IG-LIKE C2-TYPE 35.
 FT DOMAIN 4965 5057 IG-LIKE C2-TYPE 36.
 FT DOMAIN 5067 5160 IG-LIKE C2-TYPE 37.
 FT DOMAIN 5171 5260 IG-LIKE C2-TYPE 38.
 FT DOMAIN 5277 5366 IG-LIKE C2-TYPE 39.
 FT DOMAIN 5383 5472 IG-LIKE C2-TYPE 40.
 FT DOMAIN 5487 5578 IG-LIKE C2-TYPE 41.
 FT DOMAIN 5595 5685 IG-LIKE C2-TYPE 42.
 FT DOMAIN 5701 5790 IG-LIKE C2-TYPE 43.
 FT DOMAIN 5815 5904 IG-LIKE C2-TYPE 44.
 FT DOMAIN 5925 6014 IG-LIKE C2-TYPE 45.
 FT DOMAIN 6038 6130 IG-LIKE C2-TYPE 46.
 FT DOMAIN 6150 6239 IG-LIKE C2-TYPE 47.
 FT DOMAIN 6275 6368 FIBRONECTIN TYPE-III.
 FT DOMAIN 6413 6502 IG-LIKE C2-TYPE 48.
 FT DOMAIN 6507 6596 IG-LIKE C2-TYPE 49.
 FT DISULFID 568 621 POTENTIAL.
 FT DISULFID 2908 2975 POTENTIAL.
 FT DISULFID 3015 3065 POTENTIAL.
 FT DISULFID 3707 3759 POTENTIAL.
 FT DISULFID 3826 3890 POTENTIAL.
 FT DISULFID 5092 5157 POTENTIAL.
 FT DISULFID 5298 5350 POTENTIAL.
 FT DISULFID 5508 5560 POTENTIAL.
 FT DISULFID 5616 5669 POTENTIAL.
 FT DISULFID 5722 5764 POTENTIAL.
 FT DISULFID 5836 5901 POTENTIAL.

```

FT DISULFID 5946 5998 POTENTIAL.
FT DISULFID 6036 6171 POTENTIAL.
FT DISULFID 6421 6486 POTENTIAL.
FT CONFLICT 2137 2137 A -> P (IN REF. 1).
FT CONFLICT 2245 2247 A -> P (IN REF. 1).
FT CONFLICT 2258 2258 A -> P (IN REF. 1).
FT CONFLICT 2284 2284 E -> G (IN REF. 1).
FT CONFLICT 2297 2297 M -> I (IN REF. 1).
FT CONFLICT 3531 3531 A -> G (IN REF. 1).
FT CONFLICT 3884 3888 DAGEY -> RRRRI (IN REF. 1).
FT CONFLICT 3929 3929 A -> V (IN REF. 1).
FT CONFLICT 5134 5134 A -> P (IN REF. 1).
FT CONFLICT 5145 5145 T -> S (IN REF. 1).
FT CONFLICT 5185 5185 G -> A (IN REF. 1).
FT CONFLICT 5199 5199 K -> N (IN REF. 1).
FT CONFLICT 5202 5202 L -> F (IN REF. 1).
FT CONFLICT 5213 5213 F -> L (IN REF. 1).
FT CONFLICT 6178 6178 A -> G (IN REF. 1).
FT CONFLICT 6268 6268 K -> E (IN REF. 1).
SQ SEQUENCE 6632 AA; 73165 MW; 26203EDD62960B89 CRC64;

Query Match
Best Local Similarity 23.6%; Score 164; DB 1; Length 6632;
Matches 107; Conservative 59; Mismatches 175; Indels 112; Gaps 22;

QY 27 KVVLLGKGDVETLTCTAQQKSIQF-HMKNQIKILGNQSFLLTGSPKLNDRADSRSS 85
DB 3693 KETFAVEGVTVLECKVKNKESHPOIKFKNQDPVEL-----GQHM-----QKEE 3735
QY 86 LWDQGNFPLIKNLKIEDSDTYICE-----VED-----QKEE 117
DB 3736 VLEDGNIKLTIQNAKEDVGARCEAVNVAGKANTNADIKIQAFAVEHVTDESGOLEE 3795
QY 118 V-QLVFGITANSIDTHLOG-----OSLTLESPPGSSPSVQCR-----SPRGKNI 163
DB 3796 IQOFETVGTASSKTDGTGGAPEFVELRSCTVTEK-----QQAIIKCKYKGGPRPKIKWT 3851
QY 164 QGGKTLTSSV---QLELDSDSGTWTCTV--LQNKAKYFKIDYPCAPAPKSKDKT----- 213
DB 3852 KEGKEVENSARVRAHEKDDGTLTLTFDNTQADAGEYR-----CEKENEYSAMTEGPIT 3906
QY 214 ---HTCPPELLGGPSVYLPFPKPKDITLMISRTPEVTCVVVDVSHEDPEVFENNYVDGVEYH 270
DB 3907 VTLGAPKIKDGAEPDLQPVKPAVV-----TVGETAVLEGKISGKPKSVKKYKNGEBLK 3961
QY 271 NAKTTPREEQVNSTYRNVSVLTVLHODWINGKEYCKVSKNKLPAPIEKTISKANGQPRE 330
DB 3962 PEDRVKIEMLDGTGR---LTVTNAKLDDMDREYCEASNEFGDWSVTLT-----VKE 4012
QY 331 POVYTLPPSRDELTKNOVSLT-----CLVKGFPYSDIAVEMESNGQPENNYKTPPVLD 384
DB 4013 P-AQVAPGFFKELSAIQVKEKETAFKECVSGTKPD--VKWPKDGTPLKEDKRVAFEST 4068
QY 385 SDGSFPLYSKLTVDKSRW-QQGVFSCSYWHEA 416
DB 4069 DDGT-----QRLVIEDSKTDDQGN-VRIEVSND 4096

```

```

RA Seleling E., Miller J., Wilson R., Storch U.;
RT "Evolution of mouse immunoglobulin lambda genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:4681-4685(1982).
[2]
RA SEQUENCE OF 1-18 (MYELOMA PROTEIN CBPC-49).
RA Breyer R.M., Sauer R.T., Eisen H.N.;
RT "The variable region of mouse lambda-3 chains.";
RL ICN UCLA Symp. Mol. Cell. Biol. 20:105-110(1981).
[3]
RP SEQUENCE OF 10-104 (MYELOMA PROTEIN CBPC-49 AND MONOCLONAL ANTIBODY
RP 8-47).
RX MEDLINE=81223782; PubMed=6165998;
RA Azuma T., Steiner L.A., Eisen H.N.;
RT "Identification of a third type of lambda light chain in mouse
RL immunoglobulins.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:569-573(1981).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC This SMS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, J00585; AAB59670.1; -.
CC F1R; B93922; U3MS.
CC DR HSP; P01842; 2MCG.
CC DR InterPro: IPR007110; IG-1like.
CC DR InterPro: IPR003597; IG-cl.
CC DR InterPro: IPR003066; IG_MHC.
CC DR Pfam: PF00047; Ig_1.
CC DR SMART; SM00407; IGc1; 1.
CC DR PROSITE; PS50835; IG_LIKE; 1.
CC DR PROSITE; PS00290; IG_MHC; 1.
CC KW Immunoglobulin domain; Immunoglobulin C region.
CC FT NON_TER 1 1
CC FT DOMAIN 6 99 IG-LIKE.
CC FT DISULFID 27 85
CC FT DISULFID 103 103 INTERCHAIN (WITH HEAVY CHAIN).
CC SQ SEQUENCE 104 AA; 11371 MW; 83CEBCD4AA348E1 CRC64;

Query Match
Best Local Similarity 6.0%; Score 163; DB 1; Length 104;
Matches 35; Conservative 16; Mismatches 47; Indels 4; Gaps 2;

QY 327 QPRE-POVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLD 385
DB 1 QPKSTPILTMPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLD 60
QY 386 DGSFPLYSKLTVDKSRW-QQGVFSCSYWHEALHNYTKSLS 427
DB 61 DKWYASSFLHLTSPDQWRSNHSFTQVTHE---GDVTEKSL 99

```

```

RESULT 71
LAC3 MOUSE STANDARD; PRT; 104 AA.
ID LAC3 MOUSE STANDARD; PRT; 104 AA.
AC P01845;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE IG lambda-3 chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83014953; PubMed=6812053;

```

```

RESULT 72
PGBM MOUSE STANDARD; PRT; 3707 AA.
ID PGBM MOUSE STANDARD; PRT; 3707 AA.
AC Q05753;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (Perlecan) (PLC).
GN HSPG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Melanoma;

```


CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M27749; AAA36100.1; -
 DR EMBL; M34513; AAA36096.1; -
 DR EMBL; M34511; AAA36096.1; JOINED.
 DR EMBL; M34512; AAA36096.1; JOINED.
 DR EMBL; BC012293; AAH12293.1; -
 DR PIR; A33911; A33911.
 DR HSSP; P01842; 7PAB.
 DR Genew; HGNC:5870; IGLL1.
 DR MIM; 146770; -
 DR GO; GO:0016020; C:membrane; NAS.
 DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003597; Ig-cl.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00407; IGL1; 1.
 DR PROSITE; PS00290; IG_MHC; 1.
 DR Antigen; Signal; Immunoglobulin domain.
 KM SIGNAL 1 37 POTENTIAL.
 FT CHAIN 38 213 IMMUNOGLOBULIN LAMBDA-LIKE POLYPEPTIDE 1.
 FT DOMAIN 97 108 J REGION (BY SIMILARITY TO LAMBDA
 FT DOMAIN 109 213 C REGION (BY SIMILARITY TO LAMBDA
 FT DOMAIN 109 213 LIGHT-CHAIN).
 SQ SEQUENCE 213 AA; 22963 MW; 9133A7742B943C79 CRC64;

Query Match
 Best Local Similarity 30.5%; Score 162.5; DB 1; Length 213;
 Matches 36; Conservative 21; Mismatches 44; Indels 17; Gaps 3;

QY 327 QPR-EPQVYTLPPSRDELTKQVSLTCLVKGFPSPDIWVESNGQP-ENNYKTPPYLD 384
 DB 109 QPKATPSVTLFPSSSEELQANKATLVCLMNDPFGILVTWKADGTPITQGVEMTTPSKQ 168
 QY 385 SDGSFFLYSKLTVDSKRWQGNVFCSSVMEALNHHYTKSLSPGLQDDETCAEQ 442
 DB 169 SNNKYAASSYSLTPEQWRSSRSYSCQVME-----GSTVEKTVAPAR 211

RESULT 74
 LAC2_MOUSE
 ID LAC2_MOUSE STANDARD; PRT; 104 AA.
 AC P01844;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ig lambda-2 chain C region.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83014953; PubMed=6812053;
 RA Seising E., Miller J., Wilson R., Storb U.;
 RT "Evolution of mouse immunoglobulin lambda genes";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:4681-4685(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82274221; PubMed=6287422;
 RA Wu G.E., Govindji N., Hozumi N., Murialdo H.;
 RT "Nucleotide sequence of a chromosomal rearranged lambda 2
 RL Immunoglobulin gene of mouse";
 RN Nucleic Acids Res. 10:3831-3843(1982).
 RN [3]
 RP SEQUENCE FROM N.A. (MOPC 315).
 RX MEDLINE=82220143; PubMed=6283385;

RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
 RA Baltimore D.;
 RT "Somatic variants of murine immunoglobulin lambda light chains";
 RL Nature 298:380-382(1982).
 RN [4]
 RP SEQUENCE (MOPC 315).
 RX MEDLINE=74048693; PubMed=4760498;
 RA Dugan E.S., Bradshaw R.A., Simms E.S., Eisen H.N.;
 RT "Amino acid sequence of the light chain of a mouse myeloma protein
 RL (MOPC-315)";
 RL Biochemistry 12:5400-5416(1973).
 RN [5]
 RP SEQUENCE OF 66-104 (MOPC 315), AND REVISIONS.
 RX MEDLINE=81223782; PubMed=6165998;
 RA Azuma T., Steiner L.A., Eisen H.N.;
 RT "Identification of a third type of lambda light chain in mouse
 RL immunoglobulins";
 RL Proc. Natl. Acad. Sci. U.S.A. 78:569-573(1981).
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; J00595; AAA3151.1; -
 DR PIR; C93922; LZMS.
 DR HSSP; P01842; 2MCG.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003597; Ig-cl.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00407; IGL1; 1.
 DR PROSITE; PS00290; IG_LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON TER 1 1
 FT DOMAIN 6 99 IG-LIKE.
 FT DISULFID 27 85
 FT DISULFID 103 103 INTERCHAIN (WITH HEAVY CHAIN).
 FT DISULFID 103 103
 SQ SEQUENCE 104 AA; 11254 MW; CE4B67B868862D3 CRC64;

Query Match
 Best Local Similarity 34.3%; Score 161; DB 1; Length 104;
 Matches 35; Conservative 15; Mismatches 48; Indels 4; Gaps 2;

QY 327 QPRE-POVYTLPPSRDELTKQVSLTCLVKGFPSPDIWVESNGQENNYKTPPYLD 385
 DB 1 QPKATPSVTLFPSSSEELKANKATLVCLINFSFGTVAMKANGTITQGVDSINPTKE 60
 QY 386 DGSFFLYSKLTVDSKRWQGNVFCSSVMEALNHHYTKSL 427
 DB 61 GNKFMASFLHLSQWRSHNSFTQVTHE---GDVTEKSL 99

RESULT 75
 HB2D_PIG
 ID HB2D_PIG STANDARD; PRT; 258 AA.
 AC P15983;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SIA Class II histocompatibility antigen, DQ haplotype D beta chain
 DE precursor.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9623;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=90361905; PubMed=2391424;
 RA Guettersen K., Leguenn C., Hirsch F., Germana S., Pratt K.,
 RA Sachs D.H.;
 RT "Class II gene of miniature swine. IV. Characterization and
 RT expression of two allelic class II DQB cDNA clones.";
 RL J. Immunol. 145:1946-1951(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL: M31498; AAA1085.1; --
 DR HSP; P13760; 2589.
 DR InterPro: IPR007110; Ig-1like.
 DR InterPro: IPR003597; Ig_C1.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR000353; MHC_II_beta.
 DR Pfam: PF00047; Ig_1.
 DR Pfam: PF00969; MHC_II_beta; 1.
 DR ProDom: PD000328; MHC_II_beta; 1.
 DR SMART; SM00407; IgC1; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; 1.
 DR MHC II; Transmembrane; Glycoprotein; Signal.
 KM
 FT SIGNAL
 FT CHAIN 32 258
 FT 1 31
 FT DOMAIN 32 123 SLA CLASS II HISTOCOMPATIBILITY ANTIGEN,
 FT DOMAIN 124 217 DO HAPLOTYPE D BETA CHAIN.
 FT DOMAIN 218 227 EXTRACELLULAR BETA-1.
 FT TRANSMEM 228 248 CONNECTING PEPTIDE.
 FT DOMAIN 249 258 CYTOPLASMIC TAIL.
 FT DISULFID 146 108 BY SIMILARITY.
 FT CARBOHYD 48 42 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 258 AA; 29262 MW; 53AC75110AED47C3 CRC64;
 Query Match 6.0%; Score 161; DB 1; Length 258;
 Best Local Similarity 28.4%; Pred. No. 0.0015;
 Matches 46; Conservative 28; Mismatches 80; Indels 8; Gaps 4;
 QY 262 WYDGVENAKTKPREEQNSTYRVSVLTLHODMLNGEKYCKVSKALPAPIEKTI 321
 DB 54 WSVDRY-IYNOEEFLFSDDMGEYRAVTLGRPDADYLANGQKAELEQKAEIDTVCKHNY 112
 QY 322 SKAKG---QPREQVYTLPRSRDELTKQVSLTLVKGFPYPDIAVEMSNQOPEN-NYK 377
 DB 113 QIEEGTTLRRVQPTVLTISPKAEALNNHNLVCAVLTDPYPOVKQWFRNGQDETAGV 172
 QY 378 TTPPVLDSDGFPLYSKLTIVDKSRMOQGVNFCVGHALHN 419
 DB 173 STPLIRNGD---WTGYVLVLENNLQRGDVTYTRVHSSLOS 211
 RESULT 76
 VCA1 HUMAN STANDARD; PRT; 739 AA.
 AC P19320;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Vascular cell adhesion protein 1 precursor (V-CAM 1) (CD106 antigen)
 DE (INCM-100).
 GN VCA1 OR L1CAM
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RX [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=umbilical vein;
 RX MEDLINE=91016951; PubMed=1699207;
 RA Polte T., Newman W., Gopal T.V.;
 RT "Full length vascular cell adhesion molecule 1 (VCAM-1).";
 RL Nucleic Acids Res. 18:5901-5901(1990).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90090619; PubMed=268898;
 RA Osborn L., Hession C., Tizard R., Vassallo C., Luhnowskyj S.,
 RA Chi-Rosso G., Lobb R.;
 RT "Direct expression cloning of vascular cell adhesion molecule 1, a
 RT cytokine-induced endothelial protein that binds to lymphocytes.";
 RL Cell 59:1203-1211(1989).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91352090; PubMed=1715583;
 RA Cybulsky M.I., Fries J.W., Williams A.J., Sultan P., Eddy R.,
 RA Byers M., Shows T., Gimbrone M.A. Jr., Collins T.;
 RT "Gene structure, chromosomal location, and basis for alternative mRNA
 RT splicing of the human VCAM1 gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:7859-7863(1991).
 RN (4)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91201302; PubMed=1707873;
 RA Hession C., Tizard R., Vassallo C., Schiffer S.B., Gott D., Moy P.,
 RA Chi-Rosso G., Luhnowskyj S., Lobb R., Osborn L.;
 RT "Cloning of an alternate form of vascular cell adhesion molecule-1
 RT (VCAM1).";
 RL J. Biol. Chem. 266:6682-6685(1991).
 RN (5)
 RP SEQUENCE FROM N.A., AND VARIANTS PHE-318; ALA-384; ALA-413 AND
 RP LEU-716.
 RA Rieder M.J., Arnel T.Z., Carrington D.P., Osuna M., Kuldanek S.A.,
 RA Rajkumar N.R., Tsch E.J., Yi Q., Nickerson D.A.;
 RT Submitted (Aug-2002) to the EMBL/Genbank/DBJ databases.
 RN (6)
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RC TISSUE=Retinal pigment epithelium;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stabileton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.T., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (7)
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 25-226.
 RX MEDLINE=95147978; PubMed=7531291;
 RA Jones E.Y., Harlos K., Bottomley M.J., Robinson R.C., Driscoll P.C.,
 RA Edwards R.M., Clements J.M., Dudgeon T.J., Stuart D.I.;
 RT "Crystal structure of an integrin-binding fragment of vascular cell
 RT adhesion molecule-1 at 1.8-A resolution.";
 RL Nature 373:539-544(1995).
 RN (8)
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 25-220.
 RX MEDLINE=95296382; PubMed=753925;
 RA Wang J.-H., Pepinsky R.B., Stehle T., Liu J.-H., Karpusas M.,
 RA Browning B., Osborn L.;

RT "The crystal structure of an N-terminal two-domain fragment of
RT vascular cell adhesion molecule 1 (VCAM-1): a cyclic peptide based on
RT the domain 1 C-D loop can inhibit VCAM-1-alpha 4 integrin
RT interaction.";
RT Proc. Natl. Acad. Sci. U.S.A. 92:5714-5718(1995).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 25-220.
RA Wang J.-H., Stehle T., Pepinsky R.B., Liu J.-H., Karpusas M.,
RA Osborn L.;
RT "Structure of a functional fragment of VCAM-1 refined at 1.9-A
RT resolution.";
RT Acta Crystallogr. D 52:369-379(1996).
CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION
CC IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1
CC INTEGRIN VIA4 ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL
CC TRANSDUCTION. THE VCAM1/VIA4 INTERACTION MAY PLAY A
CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE
CC EMIGRATION TO SITES OF INFLAMMATION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=Long; P19320-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P19320-2; Sequence=VSP_002580;
CC -1- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as
CC well as on macrophage-like and dendritic cell types in both normal
CC and inflamed tissue.
CC -1- INDUCTION: By cytokines (e.g. IL-1, TNF-alpha).
CC -1- PTM: Sialoglycoprotein.
CC -1- DISEASE: May play an important role in the genesis of
CC atherosclerosis and rheumatoid arthritis.
CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD106 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd106.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X53051; CAA37218.1; -;
DR EMBL; M30257; AAA51917.1; ALT_TERM.
DR EMBL; M73255; AAA61270.1; -;
DR EMBL; M60335; AAA61269.1; -;
DR EMBL; AF536818; AAM96190.1; -;
DR EMBL; BC017276; AAM17276.1; -;
DR PIR; A41288; A41288.
DR PIR; B41288; B41288.
DR PDB; 1VCA; 15-SEP-95.
DR PDB; 1VSC; 20-JUN-96.
DR PDB; 1IJ9; 07-NOV-01.
DR Genew; HGNC:12663; VCAM1.
DR MIM; 192225; -;
DR InterPro; IPR003987; ICAM VCAM-1.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003989; VCAM-1.
DR Pfam; PF00047; Ig; 6.
DR PRINTS; PRO1472; ICAMVCAM1.
DR PRINTS; PRO1474; VCAM1.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS50835; IG_LIKE; 5.
KW Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;
KW Repeat; Signal; Alternative splicing; Polymorphism; 3D-structure.
FT CHAIN 1 24
FT DOMAIN 25 739 VASCULAR CELL ADHESION PROTEIN 1.
FT 698
FT TRANSMEM 699 720 EXTRACELLULAR (POTENTIAL).
FT

FT DOMAIN 721 739 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 105 IG-LIKE C2-TYPE 1.
FT DOMAIN 109 212 IG-LIKE C2-TYPE 2.
FT DOMAIN 223 309 IG-LIKE C2-TYPE 3.
FT DOMAIN 312 399 IG-LIKE C2-TYPE 4.
FT DOMAIN 408 506 IG-LIKE C2-TYPE 5.
FT DOMAIN 511 595 IG-LIKE C2-TYPE 6.
FT DOMAIN 600 684 IG-LIKE C2-TYPE 7.
FT DISULFID 47 95
FT DISULFID 52 99
FT DISULFID 137 195
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 310 402 Missing (in isoform Short).
FT VARIANT 318 318 S -> F.
FT VARIANT 384 384 /FTId=VAR_014309.
FT VARIANT 413 413 T -> A.
FT VARIANT 413 413 /FTId=VAR_014310.
FT VARIANT 716 716 G -> A.
FT VARIANT 716 716 /FTId=VAR_014311.
FT VARIANT 716 716 I -> L.
FT STRAND 26 30 /FTId=VAR_014312.
FT STRAND 34 38
FT TURN 39 40
FT STRAND 43 50
FT STRAND 56 61
FT TURN 62 63
FT STRAND 70 74
FT TURN 75 76
FT STRAND 82 82
FT STRAND 87 89
FT HELIX 92 99
FT STRAND 100 101
FT TURN 102 114
FT STRAND 120 123
Query Match 5.9%; Score 160.5; DB 1; Length 739;
Best Local Similarity 19.5%; Pred. No. 0.0062;
Matches 114; Conservative 87; Mismatches 203; Indels 181; Gaps 25;
QY 16 LALLPAAATGQNKV-----VLGKQDTELTCTASQKSIQPHWKNNOIKIG---NQ 65
DB 15 LWMFPAASQAFKIETTPESRYLAQIGDSVGLTCTCESPFPSMRQIDSPILNGKYTNE 74
QY 66 G--SFLTKGPSKLNDR-----ADSRRLMDQG-----NPF----- 93
DB 75 GTTSTLTMTNPFVSFGNEHSYLTCTATCSRK--LEKGIQVEIYSFPKDEIHLSGPLEAKP 132
QY 94 -----LIIKNLK-----TEDSD-----TYICEVED----- 113
DB 133 ITVKCSVADVYPPDRLEIDILKGDHLMKSGQPELDDNRKSLFTKSLVETFPVIEDIGKV 192
QY 114 -----QKEEVOLLVFG---LTSNDTHLLQCGSLTLTLESPPGS 149
DB 193 LVCRAKLHIDEMDSVPTVROAVKELQYIISPKNTVIVSNPSTKLOEGSGVTMTCSSEGLP 252
QY 150 SPSPV---QCRSPRGKNIQGGKTLVSQLELDQSGTWTC---TVLONOKKVEFKIDIVP 201
DB 253 APEIFMSKLDNGNLQHLNSGNATLTLLAMMEDSGIYCGVNLIGKRKREVELIVQ--- 309
QY 202 CPAPEPKSCDKTHTCCELLGGPSVFLPPPKQDTLMISRTPEYTCVVADVSHEDPEVKFN 261
DB 310 -----EK-----PFTVEISPEGRIRAAQIGDSVMTLCSVMGC--ESPSPSWR 348
QY 262 WYVDGVEVNAKTKPREEOQNSYTRVVSVLTVTHQDMLNGKEKYCKYS--NKALPAPIEK 319
DB 349 TQIDS-----PLSGKVSSEGTNST-----LTLSPVSPFENHSYLTCTVCGHKLEKGIQV 398

```

OY 330 TISKAKGOREQVYTLPPSRRELTRNQVSLCYKGFPSD-IANE----- 365
DB 399 ELYSF--PRDEPI--EMSGGLVWGSSVTVSCCKPSVPLRLRELTLEIKGETTENIEF 452
OY 366 WESNQCPENNYKT-----TPVLDSGDSFEFLSKLTVDSRMQOGGVFSCSVMEALAHN 420
DB 453 LEDTDMKSLSENSLEMTFTPTLEDTRKALVCAQKHLIDMEEPKROSTQTLVYVAVR 512
OY 421 YTKSLS---LSPGLQDDETCAEADQDGLDGLWTTDPRAALP 461
DB 513 DTTVLVSPSSILEGSSVNMTCLSQGFAPAPKILMSRQLNGEQLQF 557

RESULT 77
CAML_BRARE STANDARD; PRT; 1197 AA.
AC 090478;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neural cell adhesion molecule LI.1 (N-CAM LI.1) (Fragment).
GN NADLI.1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN (1)
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC TISSUE=Embryo;
RX MEDLINE=96155762; PubMed=868941;
RA Tongiorgi E., Bernhardt R.R., Schachner M.;
RT "Zebrafish neurons express two LI-related molecules during early
RT axonogenesis.";
RL J. Neurosci. Res. 42:547-561 (1995).
CC -1- FUNCTION: Cell adhesion molecule with an important role in the
CC development of the nervous system. Involved in neuron-neuron
CC adhesion, neurite fasciculation, outgrowth of neurites, etc. Binds
CC to axonin on neurons (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in postmitotic neurones in 16-36
CC hour embryos, including those in the brain, cranial ganglia and
CC oltic and olfactory placodes, and in all classes of spinal
CC neurones.
CC -1- DEVELOPMENTAL STAGE: Onset of expression correlates with the
CC initiation of axonogenesis in 16-36 hour embryos.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC LI/neurofascin/NG2CAM family.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 5 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (see http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, X89204; CAA61490.1; -
CC DR PIR: T30581; T30581.
CC DR HSSP: P20241; 1CPR.
CC DR ZFIN, ZDB-GENE-980526-512; nadli.1.
CC DR InterPro, IPR008957; FN_III-like.
CC DR InterPro, IPR003961; FN_III.
CC DR InterPro, IPR003962; FNIII subd.
CC DR InterPro, IPR007110; Ig_1Ike.
CC DR InterPro, IPR003598; Ig_C2.
CC DR Pfam, PF00041; fn3; 5.
CC DR Pfam, PF00047; Ig_6.
CC DR PRINTS, PR00014; FNTYPEIII.
CC DR SMART, SM00060; FN3; 5.

```

Query Match	Similarity	Score	DB	Length
Beat Local	18.4%	Pred. No. 0.012;		1197;
Matches 107;	Conservative 84;	Mismatches 208;	Indels 182;	Gaps 25;
DR	SMART: SM00408; IGC2: 4.			
DR	PROSITE: PS50835; IG LIKE; 6.			
KW	Neurogenesis; Cell adhesion; Developmental protein; Glycoprotein;			
KW	Transmembrane; Repeat; Immunoglobulin domain.			
FT	DOMAIN	1	1	
FT	TRANSMEM	<1	1054	
FT	DOMAIN	1055	1075	
FT	DOMAIN	1076	1197	
FT	DOMAIN	<1	58	
FT	DOMAIN	69	160	
FT	DOMAIN	165	263	
FT	DOMAIN	268	355	
FT	DOMAIN	360	442	
FT	DOMAIN	451	541	
FT	DOMAIN	546	638	
FT	DOMAIN	645	739	
FT	DOMAIN	744	849	
FT	DOMAIN	850	948	
FT	DOMAIN	952	1029	
FT	DISULFID	92	143	
FT	DISULFID	199	247	
FT	DISULFID	289	339	
FT	DISULFID	383	432	
FT	DISULFID	472	525	
FT	CARBOHYD	135	135	
FT	CARBOHYD	149	149	
FT	CARBOHYD	221	221	
FT	CARBOHYD	298	298	
FT	CARBOHYD	414	414	
FT	CARBOHYD	421	421	
FT	CARBOHYD	438	438	
FT	CARBOHYD	449	449	
FT	CARBOHYD	708	708	
FT	CARBOHYD	959	959	
FT	CARBOHYD	968	968	
FT	CARBOHYD	1002	1002	
FT	CARBOHYD	1027	1027	
SO	SEQUENCE	1197 AA;	132860 MW;	7CE1505EEFAC7B28 CRC64;
QY	HLHLVLQLALPLPAATQGNKVVLGKGGDTVELLTCTASQKSIQIFMKNSNQIKILGNQGSF	68		
DB	HYTATVEAA--PYMTSPSEBHLVAPGEIVRLDCKADGIPANIMW--SINGVPVSGTD---	312		
QY	69 LTKGPSKLNDRAISRSLMDGQNFPLIIKNIKIEDSDTYICEVEDQKEVQVLVGLTAN	128		
DB	313 -----VDRRRRV---SSGKLLISNVEFSDTAVYQCEAVNAKHSILDI-----N	351		
QY	129 SDTHLLQGQSILTL-----LESPPGSSSVQCRS-----PRGNIOGK	167		
DB	352 THVHVELVPAOLILPDERLYVATAGQVYMLDCKRTFGSPPLKRIHWEILDISIPALSMAKISQ	411		
QY	168 T-----LVSQLELDQSGTWTCTVYQONQKVEFKIDIVCPAPBPKSCDKTTCPELLGAP	223		
DB	412 TTNGSLKISNVSSEDSNRITYSVSEETKSIADVEVL-----NRTK-----LVGP	456		
QY	224 SVFLPPPKKDTLMISRTPEVTCVAVVDSHEDPEVKFNWYVDGVEVHNAKTKPREQYNS	283		
DB	457 FQNLHVRIGSDAILHCK-----YTVDHNLKSPVQ--MNKDG--HKITASTNDKYHE	505		
QY	284 TYRVAVSVTLVHQMWNKKEYKCVSNKALPAPLEKTIKSKAKGPREPQVYTLPPSDEL	343		
DB	506 IEGSLKVLVDYQEMDM--GIYSCVST--TLDSDTASGYITVQDKPDPQSIKLSKMER-	560		
QY	344 TKNGVSLTCLVKGFPSPDIAVENESNGQPEKNYKTPPEVL-----SDGSFPLYSKLT	396		
DB	561 -----SVTISMPSP--VENNSVTEVEVLEMGEGTTPDEQGMQKYSVS	601		
QY	397 VDKSRMOQGNVFGSCVMEHLH-----NHYYTKSLSL-----	429		

```

Db      602 QDIDSWR--SICSYKHYFOIRAVNSIGTSAPTESSLSYSTPAKPDTPNENWTLSTDP 659
Qy      430 -----PGLDLDIFCAKQDGEILGIMTTPPPRSALPAP-----TGS 467
Db      660 KSMITISQMBMRQNRQNGPQFKVFWRRRAADS--GAHWI-----SSVSNPPLMVNNTGT 712
Qy      468 -----ALPDPTQASALPPPPASALPALALAV 493
Db      713 FVSFEIKVOAVNDLGAAPPLVIVIGSGEDPFLKAPSLALSV 753

RESULT 78
LAC_HUMAN
ID      LAC_HUMAN          STANDARD;      PRT;      105 AA.
AC      P01842; P80423;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Ig lambda chain C regions.
GN      IGLC1 AND IGLC2 AND IGLC3.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE (BENCE-JONES PROTEIN SH).
RX      MEDLINE=7016723; PubMed=4909564;
RA      Titani K., Wikler M., Shinoda T., Putnam F.W.;
RT      "The amino acid sequence of a lambda type Bence-Jones protein. 3. The
RT      complete amino acid sequence and the location of the disulfide
RT      bridges."
RL      J. Biol. Chem. 245:2171-2176(1970).
RN      [2]
RP      SEQUENCE (BENCE-JONES PROTEIN X), AND DISULFIDE BONDS.
RX      MEDLINE=6908380; PubMed=4883841;
RA      Milstein C., Clegg J.B., Jarvis J.M.;
RT      "Immunoglobulin lambda-chains. The complete amino acid sequence of a
RT      Bence-Jones protein."
RL      Biochem. J. 110:631-652(1968).
RN      [3]
RP      SEQUENCE (BENCE-JONES PROTEIN NIG-64).
RX      MEDLINE=83186114; PubMed=6404900;
RA      Kametani F., Takayasu T., Suzuki S., Shinoda T., Okuyama T.,
RA      Shimizu A.;
RT      "Comparative studies on the structure of the light chains of human
RT      immunoglobulins. IV. Assignment of a subgroup."
RL      J. Biochem. 93:421-429(1983).
RN      [4]
RP      SEQUENCE (BENCE-JONES PROTEIN KERN).
RX      MEDLINE=71150336; PubMed=5549568;
RA      Ponstingl H., Hees M., Hilschmann N.;
RT      "Structural role of antibodies. Primary structure of a monoclonal
RT      immunoglobulin-L-chain of the lambda type, subgroup IV (Bence-Jones-
RT      protein Kern). V. The complete amino acid sequence and its genetic
RT      interpretation."
RL      Hoppe-Seyler's Z. Physiol. Chem. 352:247-266(1971).
RN      [5]
RP      PARTIAL SEQUENCE (MYELOMA PROTEIN NEMM).
RX      MEDLINE=74109253; PubMed=4814727;
RA      Chen B.L., Poljak R.J.;
RT      "Amino acid sequence of the (lambda) light chain of a human myeloma
RT      immunoglobulin (IGG New)."
RL      Biochemistry 13:1295-1302(1974).
RN      [6]
RP      SEQUENCE (DOT).
RX      MEDLINE=95255298; PubMed=7737190;
RA      Scoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
RT      "Characterization of the two unique human anti-flavin monoclonal
RT      immunoglobulins."
RL      Eur. J. Biochem. 228:886-893(1995).
RN      [7]
RP      X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF NEMM.
RX      MEDLINE=75046825; PubMed=4215080;

```

```

RA      Poljak R.J., Amzel L.M., Avey H.P., Chen B.L., Phizackerley R.P.,
RA      Saul P.;
RT      "The three-dimensional structure of the fab' fragment of a human
RT      myeloma immunoglobulin at 2.0-A resolution."
RL      Proc. Natl. Acad. Sci. U.S.A. 71:3440-3444(1974).
RN      [8]
RP      PARTIAL SEQUENCE (BENCE-JONES PROTEIN MCG).
RX      MEDLINE=75013804; PubMed=4415202;
RA      Felt J.W., Deutsch H.F.;
RT      "Primary structure of the Mcg lambda chain."
RL      Biochemistry 13:4102-4114(1974).
RN      [9]
RP      X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF MCG.
RA      Edmundson A.B., Ely K.R., Abola E.E., Schiffer M.,
RA      Panagiotopoulos N.;
RT      "Rational allotetramer and divergent evolution of domains in
RT      immunoglobulin light chains."
RL      Biochemistry 14:3953-3961(1975).
RN      [10]
RP      X-RAY CRYSTALLOGRAPHY OF MCG.
RX      MEDLINE=90133913; PubMed=2515285;
RA      Ely K.R., Herron J.N., Harker M., Edmundson A.B.;
RT      "Three-dimensional structure of a light chain dimer crystallized in
RT      water. Conformational flexibility of a molecule in two crystal
RT      forms."
RL      J. Mol. Biol. 210:601-615(1989).
RN      [11]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=82080680; PubMed=6273747;
RA      Hiter P.A., Hoillis G.P., Korsmeyer S.J., Waldmann T.A., Leder P.;
RT      "Clustered arrangement of immunoglobulin lambda constant region genes
RT      in man."
RL      Nature 294:536-540(1981).
RN      [12]
RP      MISCELLANEOUS: The sequence shown is the Kern-Oz-/Mcg- chain
RP      found in proteins SH, X, and NIG-64. The Kern protein has the
RP      Kern+ marker, the NEMM protein has the Oz+ marker, the Mcg protein
RP      has the Kern+ marker, and the MCG+ marker.
CC      -1- MISCELLANEOUS: Six tandem lambda-type genes were identified and
CC      the 3 most 5' were sequenced. These correspond to the Mcg sequence
CC      (lambda-1), the Kern-Oz- sequence (lambda-2) and the Kern-/Oz+
CC      sequence (lambda-3).
CC      -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC      or send an email to license@ebi.ac.uk).
CC      -----
DR      EMBL; J00253; AA59107.1; -;
DR      EMBL; L38562; AB36581.1; ALT_INIT.
DR      EMBL; X51754; CAB38569.1; ALT_INIT.
DR      EMBL; X51755; CA36049.1; -;
DR      EMBL; X51755; CA36051.1; -;
DR      PIR; A92057; L2HU.
DR      PDB; 2MCG; 15-JUL-92.
DR      PDB; 7FAB; 31-JAN-94.
DR      PDB; 1AOK; 04-FEB-98.
DR      PDB; 1LIL; 15-MAY-97.
DR      Genew; HGNC:5855; IGLC1.
DR      Genew; HGNC:5856; IGLC2.
DR      Genew; HGNC:5857; IGLC3.
DR      MIM; 147220; -;
DR      GO; GO:0003823; F:antigen binding; NAS.
DR      GO; GO:0006955; P:immune response; NAS.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR003597; Ig_c1.
DR      InterPro; IPR003006; Ig_MHC.
DR      Pfam; PF00047; Ig; 1.
DR      SMART; SM00407; IGL1; 1.
DR      PROSITE; PS50835; IG_LIKE; 1.

```

DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Bence-Jones protein;
 3D-structure.

FT	NON_TER	1	1	IG-LIKE.
FT	DOMAIN	6	100	
FT	DISULFID	27	86	
FT	DISULFID	104	104	
FT	VARIANT	5	5	INTERCHAIN (WITH HEAVY CHAIN).
FT	VARIANT	7	7	A -> N (IN MCG+ MARKER).
FT	VARIANT	7	7	/FTID=VAR_003898.
FT	VARIANT	45	45	S -> T (IN MCG+ MARKER).
FT	VARIANT	45	45	/FTID=VAR_003899.
FT	VARIANT	56	56	S -> G (IN KERN+ MARKER).
FT	VARIANT	82	82	/FTID=VAR_003900.
FT	VARIANT	82	82	T -> K (IN MCG+ MARKER).
FT	VARIANT	82	82	/FTID=VAR_003901.
FT	VARIANT	82	82	R -> K (IN OZ+ MARKER).
FT	VARIANT	82	82	/FTID=VAR_003902.
FT	STRAND	8	11	
FT	HELIX	15	19	
FT	TURN	20	21	
FT	STRAND	24	32	
FT	STRAND	38	43	
FT	TURN	44	45	
FT	STRAND	46	48	
FT	STRAND	52	59	
FT	TURN	61	62	
FT	STRAND	65	72	
FT	STRAND	75	80	
FT	HELIX	84	89	
FT	STRAND	94	99	
FT	STRAND	105	11236	MM; DCD9C7G201C13CC2 CRC64;
FT	SEQUENCE	105	11236	MM; DCD9C7G201C13CC2 CRC64;

Query Match
 Best Local Similarity 34.1%; Score 159; DB 1; Length 105;
 Matches 31; Conservative 18; Mismatches 40; Indels 2; Gaps 2;

327 OPR-EPQVYTLPPSRDELTKNOVSLCLVKGFPSDIAVWESNGOP-ENNYKTPPVLD 384
 1 QKRAPSVTLFPSSSELDQNKATLVCLSDPFGAVTAMKADSSPVKAGVETTPSKQ 60

385 SDGSPFLVSKLTIVDKSRMOQGNVFCGVNHE 415
 61 SNKVAASSYLSLTPEQWMSHRSSYSCQVTHE 91

Db

RESULT 79
 SHS1_HUMAN STANDARD; PRT; 503 AA.
 AC P78324; O00683; Q043799; Q8N517; Q8TAL8; Q9H022; Q9UDX2; Q9UIJ6;
 AC Q9YUJ9;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Protein-tyrosine phosphatase non-receptor type substrate 1 precursor
 DE (SHP substrate-1) (SHPS-1) (Inhibitory receptor SHPS-1) (Signal-
 DE regulatory protein alpha-1) (SIRP-alpha-1) (SIRP-alpha-2) (SIRP-alpha-
 DE 3) (MYD-1 antigen) (Brain Ig-like molecule with tyrosine-based
 DE activation motifs) (Blt) (Macrophage fusion receptor) (p84).
 GN PTPNS1 OR SHPS1 OR SIRP OR MYD1 OR BIR OR MFR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxId=9606;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RX MEDLINE=97223399; PubMed=90702220;
 RA Yano T., Matcokaki T., Amano K., Matsuda Y., Takahashi N., Ochi F.,
 RA Fujitaka Y., Kasuga M.;
 RA "Mouse and human SHPS-1: molecular cloning of cDNAs and chromosomal
 RT localization of genes";
 RL Biochem. Biophys. Res. Commun. 231:61-67(1997).
 RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1), POLYMORPHISM, PHOSPHORYLATION,
 N-GLYCOSYLATION, AND INTERACTIONS WITH PTPN11, PTPN6 AND GRB2.
 RC TISSUE=Placenta;
 RX MEDLINE=97215901; PubMed=9062191;
 RA Khaitonenkov A., Chen Z., Sures I., Wang H., Schilling J.,
 RA Ullrich A.;
 RT "A family of proteins that inhibit signalling through tyrosine kinase
 RT receptors";
 RL Nature 386:181-186(1997).
 RN (3)
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND VARIANTS ASP-95; LEU-96; ASN-100;
 RP ARG-107; GLY-109 AND VAL-131.
 RC TISSUE=Monocytes;
 RX MEDLINE=98143722; PubMed=9485180;
 RA Brooke G.P., Parsons K.R., Howard C.J.;
 RT "Cloning of two members of the SIRP alpha family of protein tyrosine
 RT phosphatase binding proteins in cattle that are expressed on monocytes
 RT and a subpopulation of dendritic cells and which mediate binding to
 RT CD4 T cells";
 RL Eur. J. Immunol. 28:1-11(1998).
 RN (4)
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS LEU-44; THR-52;
 RP ARG-54; ALA-57; GLY-75; ASP-95; LEU-96; ASN-100; ARG-107; GLY-109 AND
 RP VAL-131.
 RC TISSUE=Brain;
 RX MEDLINE=20053880; PubMed=10585853;
 RA Sano S.-I., Ohnishi H., Kubota M.;
 RT "Gene structure of mouse BIR/SHPS-1";
 RL Biochem. J. 344:667-675(1999).
 RN (5)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A.G., Coville G.J., Deadman R., Dhami P.D., Dunn M.R.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharasaho M.H., Leverhna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., Murray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prichaltingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.B., Sehra H.K., Showkeen R., Sims S.,
 RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A.R., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20";
 RL Nature 414:865-871(2001).
 RN (6)
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS LEU-44; THR-50;
 RP THR-52; ARG-54; ALA-57; GLY-75; ASP-95; LEU-96; ASN-100; ARG-107;
 RP GLY-109 AND VAL-131.
 RC TISSUE=Brain, Kidney, and Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Cavatit G.T., Scheetz T.E.,
 RA Brownstein M.J., Urdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Paley J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shechenko Y., Bouffard G.G.,
 RA Blaesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.",
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RL (7)
 RN FUNCTION, AND INTERACTIONS WITH FYB, SCAP2 AND PTK2B.
 RP MEDLINE=99401000; PubMed=10469599;
 RX Timms J.F., Swanson K.D., Marie-Cardine A., Raab M., Rudd C.E.,
 RA "SHS-1 is a scaffold for assembling distinct adhesion-regulated
 multi-protein complexes in macrophages.",
 RT Curr. Biol. 9:927-930 (1999).
 RL (8)
 RN PHOSPHORYLATION BY JAK2, AND INTERACTIONS WITH PTPN11 AND JAK2.
 RP MEDLINE=20462742; PubMed=10842184.
 RX Stoege M.R., Argersinger U.S., Wang H., Ullrich A., Carner-Su C.,
 RA "Negative regulation of growth hormone receptor/JAK2 signaling by
 RT signal regulatory protein alpha.",
 RT J. Biol. Chem. 275:28222-28229 (2000).
 RL (9)
 RN FUNCTION, AND INTERACTION WITH CD47.
 RP MEDLINE=21400825; PubMed=11509594.
 RX Lalouar S., Tanaka H., Demeure C., Mateo V., Rubio M., Brown E.J.,
 RA Maliszewski C., Lindberg F.P., Oldenborg A., Ullrich A.,
 RA Delepesse G., Sarfati M.,
 RT "Bidirectional negative regulation of human T and dendritic cells by
 RT CD47 and its cognate receptor signal-regulator protein-alpha:
 RT down-regulation of IL-12 responsiveness and inhibition of dendritic
 RT cell activation.",
 RL J. Immunol. 167:2547-2554 (2001).
 CC -1- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts
 CC as docking protein and induces translocation of PTPN6, PTPN11 and
 CC other binding partners from the cytosol to the plasma membrane.
 CC Supports adhesion of cerebellar neurons, neurite outgrowth and
 CC glial cell attachment. May play a key role in intracellular
 CC signaling during synaptogenesis and in synaptic function (By
 CC similarity). Involved in the negative regulation of receptor
 CC tyrosine kinase-coupled cellular responses induced by cell
 CC adhesion, growth factors or insulin. Mediates negative regulation
 CC of phagocytosis, mast cell activation and dendritic cell
 CC activation. CD47 binding prevents maturation of immature dendritic
 CC cells and inhibits cytokine production by mature dendritic cells.
 CC -1- SUBUNIT: Binds PTPN11 when tyrosine-phosphorylated, except in
 CC macrophages, where it primarily binds PTPN6. Binds GAB2 in vitro.
 CC Binds FGR (By similarity). Binds JAK2 irrespective of its
 CC phosphorylation status and forms a stable complex. Binds SCAP1
 CC and/or SCAP2. The resulting complex recruits FYB. Binds PTK2B.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=P78324-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P78324-2; Sequence=VSP_007030;
 CC Note=No experimental confirmation available;
 CC Name=3;
 CC IsoId=P78324-3; Sequence=VSP_007029;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in brain.
 CC Detected on myeloid cells, but not T cells. Detected at lower
 CC levels in heart, placenta, lung, testis, ovary, colon, liver,
 CC small intestine, prostate, spleen, kidney, skeletal muscle and
 CC pancreas.
 CC -1- PTM: N-glycosylated.
 CC -1- PTM: Phosphorylated on tyrosine residues in response to

CC stimulation with EGF, growth hormone, insulin and PDGF.
 CC Dephosphorylated by PTPN11.
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; D86043; BAA12974.1; -;
 CC DR EMBL; Y10375; CAAT1403.1; -;
 CC DR EMBL; Y11047; CAA71944.1; ALT_INIT.
 CC DR EMBL; AB023430; BAA87929.1; -;
 CC DR EMBL; AC004832; AAF19260.1; -;
 CC DR EMBL; AL034562; CAB38874.1; -;
 CC DR EMBL; AL049634; CAB46662.1; ALT_SEQ.
 CC DR EMBL; AL117335; CAC12723.1; -;
 CC DR EMBL; BC026692; AAB26692.1; -;
 CC DR EMBL; BC033092; AAB33092.1; -;
 CC DR EMBL; BC038510; AAB38510.1; -;
 CC DR PIR; JCS287; JCS287.
 CC DR Genew; HGNC:9662; PTPN61.
 CC DR MIM; 602461; -;
 CC DR GO; GO:0005886; C:plasma membrane; TAS.

Query Match 5.84; Score 156.5; DB 1; Length 503;
 Best Local Similarity 24.88; Pred. No. 0.0069;

Matches 79; Conservative 45; Mismatches 153; Indels 41; Gaps 14;

QY 136 GQSITLTLESPPGSSPVQCKSPRGKNIQGGKTLVSQLELDGSGTWCTVL-QNQKVE 194
 DB 73 GPABELLYNQKEGFPRTVTSSTKRENDPFSISITPADGTYICVKFRKSGPTE 132
 QY 195 FKIDIVCPAPEPKSCDKHTCPPELLGSPVFLFPFKKDTMTSRPEVTCVVDSHE 254
 DB 133 FK-----SGAGTELSTVAKPSAP-VVSGPARATP-----QHTVSFICESHGSPR 177
 QY 255 DPEVKFNWYDGVENNAKTK-PRER---QNVSTRVSVLVTLVHODMLNGKEYKKV 308
 DB 178 D--ITLWKFKNGLSPQTNVDVGVSVSYSHSTAKVLTREDSVHSGVT-----CEV 229
 QY 309 SNKALPA-PIEKT--ISKAKQPREPOVYTLPPSRDELTKVQVSLTCVKGFPSPDIABE 365
 DB 230 AHTVLQGDPLKGTANLSETRVPTLEV-TQQPRAE---NQNVTCVKKFTYFORQLT 285
 QY 366 WESNGQENNYKTPPYLDSDGSFFLYSKLTVDSRMQGNVFGCSVME--ALAHNYT 422
 DB 286 WLENGVNSRTETASTVTEKNKGTYNNMSWLVNVSARHVDKLTQGVHDQGPANVSKSD 345
 QY 423 QKSLSLSPGLQLDCTCAE 440
 DB 346 LK-VSAHPKEGSGNTAAE 362

RESULT 80

LAC_PIG STANDARD; PRT; 105 AA.

AC P01846;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

OS 19 lambda chain C region.

OS Sus scrofa (pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OC NCBI_TaxId=9823;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=78000254; PubMed=409425;

```

RA Novotny J., Franek F., Margolies M.N., Haber E.;
RT "Amino acid sequence of normal (microheterogeneous) porcine
RL Immunoglobulin lambda chains."
CC Biochemistry 16:3763-3772(1977).
CC -1- MISCELLANEOUS: This chain was obtained from a mixture of normal
CC Immunoglobulins.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR: A02129; L1PG.
DR HSSP: P01842; 2MCG.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003597; IG_c1.
DR InterPro: IPR003006; IG_MHC.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00407; IGc1; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
DR PROSITE: PS00290; IG_MHC; 1.
DR Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 100
FT DOMAIN 2 86
FT DISULFID 27 86
FT DISULFID 104 104
FT DISULFID 105 AA; 11003 MW; 3817AEBD747C396 CRC64;
SQ SEQUENCE 105 AA; 11003 MW; 3817AEBD747C396 CRC64;

Query Match
Best Local Similarity 5.8%; Score 156; DB 1; Length 105;
Matches 32; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

OY 327 QPR-EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQP-ENNYKTTTPVLVD 384
DB 1 QPKAPPLVTLFPPSSEELTKNQVSLTCLVKGFYPSDIAVEMESNQP-ENNYKTTTPVLVD 60
OY 385 SDGSFFLYSKLTVDKSRWQGQNVFSCSVWHE 415
DB 61 SNKRYAASSYTLALSADWKSQSGFTCQVTHE 91

RESULT 81
LAC5_MUSP STRAND: PRT; 105 AA.
AC P20765;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig lambda-5 chain C region.
OS Mus spretus (Western wild mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OX NCBI_Taxid=10096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88196070; PubMed=3129289;
RA Mam J., Cazenave P.A., Kindt T.J.;
RT "Conservation of the immunoglobulin C lambda 5 gene in the Mus gene."
RL EMBO J. 7:117-122(1988).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M35582; AAA39152.1; -.
CC HSSP: P01842; 2MCG.
CC InterPro: IPR007110; IG-like.
CC InterPro: IPR003597; IG_c1.
CC InterPro: IPR003006; IG_MHC.
CC Pfam: PF00047; Ig; 1.
CC SMART: SM00407; IGc1; 1.
CC PROSITE: PS50835; IG_LIKE; 1.
CC PROSITE: PS00290; IG_MHC; 1.
DR PROSITE: PS00290; IG_MHC; 1.

```

```

KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 100
FT DOMAIN 6 86
FT DISULFID 27 86
FT DISULFID 104 104
FT DISULFID 105 AA; 11674 MW; AAB417DF68471A17 CRC64;
SQ SEQUENCE 105 AA; 11674 MW; AAB417DF68471A17 CRC64;

Query Match
Best Local Similarity 5.7%; Score 153.5; DB 1; Length 105;
Matches 35; Conservative 20; Mismatches 43; Indels 5; Gaps 3;

OY 327 QPR-EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQP-ENNYKTTTPVLVD 384
DB 1 QPKSDPLVTLFPLSLKLNQANKVTLVCLVSEFPEGLTVDMKVDGVPVTVQGVETTPSKQ 60
OY 385 SDGSFFLYSKLTVDKSRWQGQNVFSCSVWHEALHNHYTKSLIS 427
DB 61 TNKRYMVSSYTLTLISDQWMPHSRYSRVYTHE---GNTVEKSVS 100

RESULT 82
LAC_RABIT STRAND: PRT; 105 AA.
AC P01847;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig lambda chain C region.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=82091105; PubMed=6797414;
RA Garcia I., Jaton J.-C.;
RT "The primary structure of the constant region of Basilea-rabbit
RT Immunoglobulin lambda-chains."
RL Biochem. J. 197:177-183(1981).
CC -1- MISCELLANEOUS: This lambda chain expresses the c7 allotypic
CC specificity.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR: A02130; L7RB.
DR HSSP: P01842; 7PAB.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003597; IG_c1.
DR InterPro: IPR003006; IG_MHC.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00407; IGc1; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
DR PROSITE: PS00290; IG_MHC; FALSE_NEG.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 100
FT DOMAIN 6 86
FT DISULFID 27 86
FT DISULFID 104 104
FT DISULFID 105 AA; 11484 MW; B427513272E8663D CRC64;
SQ SEQUENCE 105 AA; 11484 MW; B427513272E8663D CRC64;

Query Match
Best Local Similarity 5.7%; Score 153; DB 1; Length 105;
Matches 32; Conservative 17; Mismatches 45; Indels 4; Gaps 2;

OY 331 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQP-ENNYKTTTPVLVDGSGF 389
DB 6 PSVILFPPSSEELTKNQVSLTCLVKGFYPSDIAVEMESNQP-ENNYKTTTPVLVDGSGF 65
OY 390 FLYSKLTVDKSRWQGQNVFSCSVWHEALHNHYTKSLIS 427
DB 66 AASSFLHLTLNQWKSQSVTCQVTHE---GHTVEKSLA 100

RESULT 83
KACB_RABIT

```

```

ID KACB_RABIT STANDARD; PRT; 106 AA.
AC P01839;
DT 21-JUN-1986 (Rel. 01, Created)
DT 21-JUN-1986 (Rel. 01, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE 19 Kappa-B4 chain C region.
GN K-BAS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Basilea;
RA Heidmann O., Roujeon F.;
RT "Multiplicity of constant kappa light chain genes in the rabbit
RL genome: a b4b4 homozygous rabbit contains a kappa-bas gene.";
EMBL J. 2:437-441 (1983).
CC -1- MISCELANEOUS: In Basilea rabbits, the major type of light chain
CC is lambda. The kappa chain shown is a minor component. All other
CC rabbit B allotypes have Cys-64.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL/ V01241; CAA24558.1;
DR EMBL/ V00885; -; NOT_ANNOTATED_CDS.
DR PIR: A02121; K4RBBS.
DR HSSP: P01842; 72AB.
DR InterPro: IPR007110; IG_1like.
DR InterPro: IPR003597; IG_c1.
DR InterPro: IPR003006; IG_MHC.
DR Pfam: PF00047; IG_1.
DR SMART: SM00407; IGC1; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
DR PROSITE: PS00290; IG_MHC; FALSE NEG.
DR Immunoglobulin domain; Immunoglobulin C region.
FT DOMAIN 1 1
FT NON_TER 1 1
FT DISULFD 27 99 IG-LIKE.
FT DISULFD 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
SQ SEQUENCE 106 AA; 11279 MW; AF9B928DDA853849 CRC64;
Query Match 5.7%; Score 153; DB 1; Length 106;
Best Local Similarity 33.7%; Pred. No. 0.0016;
Matches 31; Conservative 19; Mismatches 36; Indels 6; Gaps 2;
Oy 326 GQPEPPQYVTPSPRDELTKQVSLTCLVKGFPYSDIAVEMESGQPE---NNYKTPP 381
Db 1 GDPAPSPVLTLPPEPSKEELTGTATIVCVANKFYSDITVTKVGTQQSGIENSKT--P 58
Oy 382 VLDSGSPFLYSKLTVDKSRWQGNVFCSCVM 413
Db 59 QSPEDNTYLSSTJLSIAQYNHSHSVYCEVV 90

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90258861; PubMed=2342461;
RA Schreye H., Thompson J., Bona M., Hefta L.J.F., Maruya A.,
RA Haasler M., Shively J.E., von Kleist S., Zimmermann W.;
RT "Cloning of the complete gene for carcinoembryonic antigen: analysis
RT of its promoter indicates a region conveying cell type-specific
RT expression.";
RL Mol. Cell. Biol. 10:2738-2748 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86038876; PubMed=3670312;
RA Beauchemin N., Benchimol S., Cournoyer D., Fuks A., Stanners C.P.;
RT "Isolation and characterization of full-length functional cDNA clones
RT for human carcinoembryonic antigen.";
RL Mol. Cell. Biol. 7:3221-3230 (1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89122014; PubMed=3220478;
RA Barnett T., Goebel S.J., Notdurft M.A., Elting J.J.;
RT "Carcinoembryonic antigen family: characterization of cDNAs coding
RT for NCA and CEA and suggestion of nonrandom sequence variation in
RT their conserved loop-domains.";
RL Genomics 3:59-66 (1988).
RN [4]
RP SEQUENCE OF 5-702 FROM N.A.
RX MEDLINE=87128144; PubMed=3814146;
RA Okkawa S., Nakazato H., Kosaki G.;
RT "Primary structure of human carcinoembryonic antigen (CEA) deduced
RT from cDNA sequence.";
RL Biochem. Biophys. Res. Commun. 142:511-518 (1987).
RN [5]
RP SEQUENCE OF 331-702 FROM N.A.
RX MEDLINE=87204247; PubMed=3033671;
RA Zimmermann W., Ortlieb B., Friedrich R., von Kleist S.;
RT "Isolation and characterization of cDNA clones encoding the human
RT carcinoembryonic antigen reveal a highly conserved repeating
RT structure.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2960-2964 (1987).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: Found in adenocarcinomas of endoderally
CC derived digestive system epithelium and fetal colon.
CC -1- PTM: COMPLEX IMMUNOREACTIVE GLYCOPROTEIN WITH A MW OF 180 KDA
CC COMPRISING 60% CARBOHYDRATE.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. CEA family.
CC -1- SIMILARITY: Contains 7 immunoglobulin-like domains.
CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD66e entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd66e.htm.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL/ M17303; AAB59513.1; -;
DR EMBL/ M59262; AAA62835.1; ALT SEQ.
DR EMBL/ M59255; AAA62835.1; JOINED.
DR EMBL/ M59256; AAA62835.1; JOINED.
DR EMBL/ M59257; AAA62835.1; JOINED.
DR EMBL/ M59258; AAA62835.1; JOINED.
DR EMBL/ M59259; AAA62835.1; JOINED.
DR EMBL/ M59260; AAA62835.1; JOINED.
DR EMBL/ M59261; AAA62835.1; JOINED.
DR EMBL/ M59709; -; NOT_ANNOTATED_CDS.
DR EMBL/ M59710; -; NOT_ANNOTATED_CDS.
DR EMBL/ M29540; AAA51967.1; -;

```



```

DR EMBL, X16455; CAA34474.1; -
DR EMBL, M15042; AAA51963.1; -
DR EMBL, M16234; AAA51972.1; -
DR PIR, A36319; A36319.
DR PDB, 1B07; 04-JUL-00.
DR Genew; HGNC:1817; CEACAMS.
DR MIM, 114890; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG_6.
DR PROSITE; PS50835; IG_LIKE; 6.
KW Immunoglobulin domain; Glycoprotein; GPI-anchor;
Membrane; Signal; Repeat; 3D-structure.
FT SIGNAL 1 34
FT CHAIN 35 685
FT PROPEP 666 702
FT DOMAIN 35 144
FT DOMAIN 146 237
FT DOMAIN 238 332
FT DOMAIN 324 415
FT DOMAIN 416 593
FT DOMAIN 594 677
FT LIPID 685 685
FT CARBOHYD 104 104
FT CARBOHYD 115 115
FT CARBOHYD 152 152
FT CARBOHYD 182 182
FT CARBOHYD 197 197
FT CARBOHYD 204 204
FT CARBOHYD 208 208
FT CARBOHYD 246 246
FT CARBOHYD 256 256
FT CARBOHYD 274 274
FT CARBOHYD 288 288
FT CARBOHYD 292 292
FT CARBOHYD 309 309
FT CARBOHYD 330 330
FT CARBOHYD 351 351
FT CARBOHYD 360 360
FT CARBOHYD 375 375
FT CARBOHYD 432 432
FT CARBOHYD 466 466
FT CARBOHYD 480 480
FT CARBOHYD 508 508
FT CARBOHYD 529 529
FT CARBOHYD 553 553
FT CARBOHYD 560 560
FT CARBOHYD 580 580
FT CARBOHYD 612 612
FT CARBOHYD 650 650
FT CARBOHYD 665 665
FT CONFLICT 320 320
SQ SEQUENCE 702 AA; 76795 MW; 6299AE26CDDDB5C CRC64;

Query Match
Best Local Similarity 18.0%; Score 151.5; DB 1; Length 702;
Matches 103; Conservative 78; Mismatches 175; Indels 217; Gaps 24;

5.6%; DB 1; Length 702;
Pred. No. 0.022;
Mismatches 175; Indels 217; Gaps 24;

32 KKGDTVELCTASQKKSIGFHKNSNQIKILGNSQFLFKGSKLNDRAISRSLMDQDN 91
250 RSGENLNLSCHASNPAPQYSFVN-----GTF-----QOST 281
92 PFLIKNLKIEDSDYICEVEDQKEVQLVFLGANSPTHLGQSLTLTLESPP----- 147
282 QSLFIPNITVNNSGSYTQAH-----NSTDTGLNRTVTITTYAEPKPF 326
148 ----GSSP-----SVQR-----SPGKNIQSGKTLVSQLEIQ 177
327 ITSNNSNPVEDDAVALTCEPEIONTTYLMMVNNQSLPVSPRLQSLNDRRTLTLSVTN 386
178 DSGTGTCTYLQNKQKVEFKIDIVPCAPAPKSCDKHTKCPRLGLGGSVFLFPPKRDITM 237

```

```

DB 367 DVGPEECGI-QNELSV-----HSDPVL---NVLGSDDE--TIS 421
QY 238 ISRT---PEVTCVVVDVSHEDPEVKFNRYVDG-VEVHNAKTKPREEQINTSTRYVSLTV 293
DB 422 PSYTYRRGVNLSLSCHASNPAPQYSWLDIGNIOOHTOE-----LPI 464
QY 294 LHQDLNKEVCKSKNKA-----LPIEKTISKAKGPREPOVYITLP 337
DB 465 SNTKNSGLYTCQANNBSGHSRTTVKTTVSAELPRP---SISSNSKPYE----- 514
QY 338 PSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGOPENNYKTPPYLDSGSPFLYSKLTV 397
DB 515 -----DDDAVAFTEPEP-----QNTTYLMMWNGQ---SLPVSPLQLSNGN-----RTLTL 558
QY 398 DKSRRQGVNFSQSVMEALNHN-----YTKSLSLSP-----GLQDLETCAEA 441
DB 559 FNVTRNDARAAYVCGIQNSVSNRSDPVLTVLGYGDTPTIISPDSYLSGANLNLSCSHA 618
QY 442 QDGLDGLMTTDPAPASALPAPPT-----GSALPDPT 474
DB 619 SNPSQYGSN-----RINGIPQOHTVLFIAKTPNNNGTYACFVSNLATGRNNSIVKSIT 673
QY 475 ASALPDPPAPASALPALAVISFLGLGVACV 507
DB 674 VASGTSGLS-----AGATVGIMIGLVGVALI 702

RESULT 85
KACA RAT
ID KACA RAT STANDARD; PRT; 106 AA.
AC P01836;
DT 21-JUL-1986 (rel. 01, Created)
DT 21-JUL-1986 (rel. 01, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Ig kappa chain C region, A allele.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA;
RX MEDLINE=82082587; PubMed=6273908;
RA Sheppard H.W., Gutman G.A.;
RT "Allelic forms of rat kappa chain genes: evidence for strong
RL selection at the level of nucleotide sequence."
RL Proc. Natl. Acad. Sci. U.S.A. 78:7064-7068(1981).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02118; K1RTA.
DR HSSP; P01842; 2MCG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00407; IGC1.1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON TER 1 1
FT DOMAIN 5 102
FT DISULFID 26 86
FT DISULFID 106 106
SQ SEQUENCE 106 AA; 11732 MW; B7B120D9700DD66 CRC64;

Query Match
Best Local Similarity 31.5%; Score 150; DB 1; Length 106;
Matches 29; Conservative 23; Mismatches 28; Indels 12; Gaps 2;

331 POVYTLPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGOPENNYKTPPYLDS----- 385
5 PTVSIFPSPMEQLTSGATVVCVANNFYPRDISVGMKIDGSGRD-----GVLDSTVTD 59

```

Qy 386 --DGSFPLYSKLTVDSKRWQGNVFCSCVMEH 415
 Db 60 SKDSTYSMSSTLSLKVEYERHNLVTCVHK 91

RESULT 86
 HB2C_PIG STANDARD; PRT; 261 AA.
 AC P15982;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SLA class II histocompatibility antigen, DQ haplotype C beta chain precursor.
 OS Sus scrofa (Pig).
 CC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 CC NCBI_TaxID=9823;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=90361905; PubMed=2391424;
 CC Guerafason K., Leguern C., Hirsch F., Germana S., Pratt K., Sachs D.H.;
 CC "Class II genes of miniature swine. IV. Characterization and expression of two allelic class II DQ beta cDNA clones.";
 CC J. Immunol. 145:1946-1951(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M31497; AAA31084.1; -
 CC EMBL; M32117; AAA53110.1; -
 CC PIR; A60404; A60404.
 CC HSSP; P13760; 2SER.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003597; Ig-cl.
 CC InterPro; IPR003006; Ig_MHC.
 CC InterPro; IPR000353; MHC_II_beta.
 CC Pfam; PF00047; Ig_1.
 CC Pfam; PF00969; MHC_II_beta_1.
 CC ProDom; PD000328; MHC_II_beta_1.
 CC SMART; SM00407; IGc1_1.
 CC PROSITE; PS50835; IG_LIKE; 1.
 CC PROSITE; PS00290; IG_MHC; 1.
 CC MHC II; Transmembrane; Glycoprotein; Signal.
 CC KW SIGNAL 1 31
 CC FT CHAIN 32 261 SLA CLASS II HISTOCOMPATIBILITY ANTIGEN,
 CC FT DOMAIN 32 261 DQ HAPLOTYPE C BETA CHAIN.
 CC FT DOMAIN 127 220 EXTRACELLULAR BETA-1.
 CC FT DOMAIN 221 230 EXTRACELLULAR BETA-2.
 CC FT TRANSMEM 231 251 CONNECTING PEPTIDE.
 CC FT DOMAIN 252 261 CYTOPLASMIC TAIL.
 CC FT DISULFID 47 111 BY SIMILARITY.
 CC FT DISULFID 149 205 BY SIMILARITY.
 CC FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 261 AA; 29550 MW; AA9581F2A3B1969D CRC64;

Query Match 5.5%; Score 149.5; DB 1; Length 261;
 Best Local Similarity 23.2%; Pred. No. 0.0085;
 Matches 56; Conservative 40; Mismatches 106; Indels 39; Gaps 8;

Qy 180 GTWCTCTVONOKKVEFKIDIVPCPAPKSCDKHTTCPELLGSPVFLPPKPKDTLMIS 239
 Db 12 GLMTAL-----TMTLVLGAPVAEGR-----SPDFFYQK-GEQCYRN 51
 Qy 240 RPEVTCVVVDVSHEDPEVKFNWYDSGVEHNNAKTKPREEQYNSTYRVVSVLTVLDHOML 299

Db 52 GTCRVRGARYINQOEHLRDESDVGFRATVPLGRPADSWNSQKVLQGRRA-EVDRV 110
 Qy 300 NGKEYCKCKVSKALPAPIEKTISKAKGPRRPQYTLPPSDELTKNQVSLTCLVKGP 359
 Db 111 CKENYQIE-----EGTTLQRRVQ-----TVTISPSAEALNHNLLVCAVTDTP 156

Qy 360 SDIAVEWESNQGPN-NYKTPPYLDSGFFLYSKLTVDSKRWQGNVFCSCVMEH 418
 Db 157 SQVAVQWFRNGQETAGVSTPLRNGD---WTQVLMLEMNIGRGVYTCRVHSSLG 213

Qy 419 N 419
 Db 214 N 214

RESULT 87
 KIL0_RAT STANDARD; PRT; 348 AA.
 ID KIL0_RAT
 AC Q920J8;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Kilon protein precursor (Kindred of IgLON).
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 CC [1]
 CC SEQUENCE FROM N.A. AND SEQUENCE OF 32-62.
 CC MEDLINE=99175207; PubMed=10075727;
 CC Funatsu N., Miyata S., Kumanooh H., Shigeta M., Hamada K., Endo Y., Sokawa Y., Maekawa S.;
 CC "Characterization of a novel rat brain glycosylphosphatidylinositol-anchored protein (Kilon), a member of the IgLON cell adhesion molecule family.";
 CC J. Biol. Chem. 274:8224-8230(1999).
 CC -1- FUNCTION: CELL-ADHESION (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain.
 CC -1- PM: Glycosylated.
 CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IgLON family.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB017139; BAA75649.1; -
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC Pfam; PF00047; Ig_3.
 CC SMART; SM00408; IGc2_2.
 CC PROSITE; PS50835; IG_LIKE; 3.
 CC Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
 CC KW Repeat; Signal.
 CC FT SIGNAL 1 31
 CC FT CHAIN 32 348
 CC FT PROPEP ? 348
 CC FT DOMAIN 32 128 REMOVED IN MATURE FORM (POTENTIAL).
 CC FT DOMAIN 133 215 IG-LIKE C2-TYPE 1.
 CC FT DOMAIN 219 307 IG-LIKE C2-TYPE 2.
 CC FT DISULFID 54 112 IG-LIKE C2-TYPE 3.
 CC FT DISULFID 154 197 POTENTIAL.
 CC FT DISULFID 239 291 POTENTIAL.
 CC FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 348 AA: 37858 MM; 37890D1CTD24ACAB CRC64;

Query Match 5.5%; Score 149; DB 1; Length 348;
 Best Local Similarity 22.4%; Pred. No. 0.013;
 Matches 72; Conservative 42; Mismatches 119; Indels 88; Gaps 14;

20 PAATGNGVYLGKGDVLTCTASQKSIQFHMKNKSIQKLGSGFLTKGP-----SK 75
 35 PMAVDNMLV--RKGDVAVRCYLEDGAS-KGAMLNRSIIIPAG--GDKMSVDPKVSIST 89
 76 LNDRADSRRLSDQGNFPLIKLKIEDSDTYICEVEDQKEVQLVFLGTLA----- 127
 90 LNKR-----DYSGLQIONVDVTDGPTCSVQCHPRMQVH-LTVQVPRKTYD 137
 128 -NSDTHLQGGSLVLTLESPPSSPSVOCR--SPRGKNIQGGKTLVSQLELDGSGTWC 184
 138 ISNDWTINEGTVTLTCLATGKPEPAISMRHISPSAKPFENGQYLDIVGTRDQGEYEC 197
 185 TV-----LQNKQVFKDIDVCPAPERKSCDKHTCELLGSPVFLPRPKDTLMIS 239
 198 SAENDVSPPDVKKVAVVNFAP-----TIGELKSG-----TTPPG 232
 240 RTEVTGVVVDVSHEDPEVKENY-----VDGEVHNAAKTRPEEQNSTYRVVSV 290
 233 RSGLRRCGAGV---PPAPFMYKGEKRLFRNGGGIITIQNSTR-----SI 275
 291 LTVLHQMNLNCKEYCKVSNK 311
 276 LTVTVTQEHFENYCVANAK 296

RESULT 88
 KAC_HUMAN STANDARD; PRT; 106 AA.
 AC P01834;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Ig kappa chain C region.
 GN IGKC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 CX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE (MYELOMA PROTEIN EU).
 RX MEDLINE=71064023; PubMed=5489770;
 RA Gortlieb P.D., Cunningham B.A., Rutishauser U., Edelman G.M.;
 RT "The covalent structure of a human gamma G-immunoglobulin. VI. Amino
 acid sequence of the light chain.";
 RL Biochemistry 9:3155-3161(1970).
 RN [2]
 RP DISULFIDE BONDS.
 RX MEDLINE=71064027; PubMed=4923144;
 RA Gall W.E., Edelman G.M.;
 RT "The covalent structure of a human gamma G-immunoglobulin. X.
 Interchain disulfide bonds.";
 RL Biochemistry 9:3188-3196(1970).
 RN [3]
 RP SEQUENCE (BENCE-JONES PROTEIN TI).
 RX MEDLINE=72188439; PubMed=5027703;
 RA Suter L., Barnikol H.U., Watanabe S., Hilschmann N.;
 RT "Rule of antibody structure. The primary structure of a monoclonal
 immunoglobulin L-chain of kappa-type, subgroup 3 (Bence-Jones protein
 TI). IV. The complete amino acid sequence and its significance for
 the mechanism of antibody production.";
 RL Hope-Seyler's Z. Physiol. Chem. 353:189-208(1972).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81042304; PubMed=6775818;
 RA Hieer P.A., Max E.E., Seldman J.G., Matzel J.V. Jr., Leder P.;

RT "Cloned human and mouse kappa immunoglobulin constant and J region
 genes conserve homology in functional segments.";
 RT Cell 22:197-207(1980).
 RN [5]
 RP SEQUENCE (BENCE-JONES PROTEIN ROY).
 RA Hilschmann N., Barnikol H.U., Hees M., Langer B., Ponstingl H.,
 RA Steinmetz-Kayne M., Suter L., Watanabe S.;
 RL (In) Franek F., Shugar D. (eds.);
 RL Gamma globulins: structure and function, pp.57-74, Academic Press,
 New York (1969).
 RN [6]
 RP SEQUENCE (BENCE-JONES PROTEIN CUM).
 RX MEDLINE=68242259; PubMed=5586923;
 RA Hilschmann N.;
 RT "The complete amino acid sequence of Bence Jones protein Cum (kappa-
 type).";
 RL Hope-Seyler's Z. Physiol. Chem. 348:1718-1722(1967).
 RN [7]
 RP SEQUENCE (BENCE-JONES PROTEIN AG).
 RX MEDLINE=69234734; PubMed=4893682;
 RA Titani K., Shinoda T., Putnam F.W.;
 RT "The amino acid sequence of a kappa type Bence-Jones protein. 3. The
 complete sequence and the location of the disulfide bridges.";
 RL J. Biol. Chem. 244:3550-3560(1969).
 RN [8]
 RP SEQUENCE (WALDENSTROM'S MACROGLOBULIN OU).
 RX MEDLINE=70201507; PubMed=5447531;
 RA Kohler H., Shimizu A., Paul C., Putnam F.W.;
 RT "Macroglobulin structure: variable sequence of light and heavy
 chains.";
 RL Science 169:56-59(1970).
 RN [9]
 RP SEQUENCE OF 1-33; 38-41 AND 62-80.
 RC TISSUE=Abdominal adipose tissue;
 RX MEDLINE=98249779; PubMed=9588180;
 RA Olsen K.E., Sletten K., Westermarck P.;
 RT "Extended analysis of AL-amyloid protein from abdominal wall
 subcutaneous fat biopsy: kappa IV immunoglobulin light chain.";
 RL Biochem. Biophys. Res. Commun. 245:713-716(1998).
 CC [1- MISCELLANEOUS: The EU sequence has the INV (3) allotypic marker,
 45-Ala and 83-Val. The ROY sequence has the INV (1,2) allotypic
 marker, 45-Ala and 83-Leu.
 CC [1- SIMILARITY: Contains 1 immunoglobulin-like domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; J00241; AAA58989.1; -;
 DR EMBL; V00557; CAA23823.1; -;
 DR PIR; B90562; K3HD.
 DR PDB; 1D5B; 09-FEB-00.
 DR PDB; 1D5T; 09-FEB-00.
 DR PDB; 1D6V; 04-OCT-00.
 DR PDB; 1HEZ; 18-JUL-03.
 DR PDB; 1HKL; 12-MAR-97.
 DR PDB; 1I7Z; 08-AUG-01.
 DR PDB; 1MIM; 15-MAY-97.
 DR Genew; HGNC:5716; IGKC.
 DR MIM; 147200; -;
 DR GO; GO:0003823; F:antigen binding; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR007110; IG-1ike.
 DR InterPro; IPR003597; IG-cl.
 DR InterPro; IPR003006; IG_MHC.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00407; IGcl; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; 1.

KW	Immunoglobulin domain; Immunoglobulin C region; 3D-structure.
FT	NON_TER 1 1
FT	DOMAIN 5 102 IG-LIKE.
FT	DISULFID 26 86 INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID 106 106 V -> L (IN INV(1,2) MARKER).
FT	VARIANT 83 83 /FTID=VAR_003897.
FT	CONFLICT 14 14 D -> N (IN REF. 7 AND 8).
FT	CONFLICT 57 57 E -> Q (IN REF. 5 AND 6).
SO	SEQUENCE 106 AA; 11609 MW; 51984BD1FDD372CE8 CRC64;
Query Match	5.5%; Score 148; DB 1; Length 106;
Best Local Similarity	29.5%; Pred. No. 0.0034;
Matches	28; Conservative 24; Mismatches 41; Indels 2; Gaps 1
Qy	331 PQVYTLPSRDELTKNOVSLTCLVGFYPSIDIAVEMESNG--QPENNYKTPPVLDSDGS 388
Db	5 PSVFIFPSPDEQLKSGTASVVCLLNFPFRKAKVQMKVDNALQSNQSGSVTEQDSKOST 64
Qy	389 FFLYSKLTVDKSRWQGNVFCSCVNHEDLNHYTQ 423
Db	65 YLSSTLTLSKADYEKHKVYACEVTHQGLSPFTK 99
RESULT: 89	
NPAS_CHICK	STANDARD; PRT; 1369 AA.
AC	042414-1; 090924;
DT	15-MAR-2004 (Rel. 43, Created)
DT	15-MAR-2004 (Rel. 43, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	Neurofascin precursor.
GN	NPASC.
OS	Gallus gallus (Chicken).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Archaeoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC	Gallus.
OX	NCBI_TaxID=9031;
RN	[1]
RP	SEQUENCE FROM N.A. (ISOFORM 2), SEQUENCE OF 26-46; 637-641; 717-730;
RP	758-781 AND 801-815, CLEAVAGE AT ARG-636, AND GLYCOSYLATION.
RC	TISSUE=Brain;
EX	MEDLINE=92317154; PubMed=1377686;
RA	Volkmmer H., Haesel B., Wolff J.M., Frank R., Rathjen F.G.;
RT	"Structure of the axonal surface recognition molecule neurofascin and
RT	its relationship to a neural subgroup of the immunoglobulin
RL	J. Cell Biol. 118:149-161(1992).
RN	[2]
RP	SEQUENCE FROM N.A. (ISOFORM 1).
RC	TISSUE=Liver;
RX	MEDLINE=98019255; PubMed=9353344;
RA	Haesel B., Rathjen F.G., Volkmmer H.;
RT	"Organization of the neurofascin gene and analysis of developmentally
RT	regulated alternative splicing."
RL	J. Biol. Chem. 272:28742-28749(1997).
CC	-1- FUNCTION: Cell adhesion, ankyrin-binding protein which may be
CC	involved in neurite extension, axonal guidance, synaptogenesis,
CC	myelination and neuron-glia cell interactions. (By similarity).
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1- ALTERNATIVE PRODUCTS:
CC	Event=Alternative splicing; Named isoforms=2;
CC	Comment=A number of isoforms are produced;
CC	Name=1;
CC	Isoid=042414-1; Sequence=Displayed;
CC	Name=2;
CC	Isoid=042414-2; Sequence=VSP_008935, VSP_008936;
CC	-1- DEVELOPMENTAL STAGE: There is one major 'early' isoform and
CC	multiple 'late' isoforms. Around 50 isoforms are found at
CC	different developmental stages.
CC	-1- PTM: N-glycosylated and O-glycosylated.
CC	-1- PTM: May be proteolytically cleaved at Arg-636.
CC	-1- SIMILARITY: Belongs to the immunoglobulin superfamily.

[illegible]

```

FT CARBOHYD 853 853 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 994 994 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1009 1009 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1133 1133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1150 1150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1171 1171 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 987 993 SCLSPV -> F (in isoform 2).
FT VARSPLIC 1132 1222 /FTid=VSP 008935.
FT VARSPLIC 1132 1222 /FTid=VSP 008936.
FT CONFLICT 1105 1105 T -> A (in REF. 2).
SQ SEQUENCE 1369 AA; 152954 MW; 835F27D086B2BF6 CRC64;

Query Match 5.5%; Score 148; DB 1; Length 1369;
Best Local Similarity 22.3%; Pred. No. 0.085;
Matches 107; Conservative 63; Mismatches 182; Indels 128; Gaps 31;

4 GVPF-----RHLLVLQLAL-----LP-----AATQGNKVVVGKKQDTVELTCTASOKKS 48
12 GIAFLALCLHLLISAIEVPLDSNIQSELPQPTITQSVVDYIVDRDNIPIECCKGNPV 71
49 IOFHKNSNQIKILGNQSF--LTGPKSLNDRADSRSLMD-----QGNFPLII 96
72 PTFSWTRN-----GKFFNVAKDP-KVSWRRRSGLTVIDFHGCGRPDDEGEYQCPA 121
97 KN-----LKIEDSDYICEVDQKEVQVLLVGLTANSSTHLLQGSLLTLLES 145
122 RNDVGTALSSKIHLOVSRSPLM-----PKEKVDYI-----EVDGAPSLQCNRP 165
146 PGSSPSV-----QCRSP--RGKNIQGGKT--LSVSOLELQDSGT-WTCT-----VLQ 188
166 PFGLEPPPIIFMSSMERPIHODKRVSOQNGDLVSNWMLQDAQIDGCMARFHTHTIQ 225
189 NOKKVEFKIDYPCPAPEPKSCDKHTTCELLIG-----PSVLEPPPKDTLMSRPE 243
226 OKNPYTLK---VKTKKPHNETSLRNHTDMYSARGTETTPS-FMYPGTSSSQMLRGVD 281
244 --VTCVVVDVSHEDDEVKFNWVDGVEVNAKTKREGQNSTYRVVSLVTLVHDDMLNG 301
282 LLLLECTIAGV---PAPDIMMYKKGGELPAGTK--LENFNALALISNVSEB-----DS 329
302 KEYKCVSNKALPARIKTIIS-KAKQGP---REPOVTLTPSRDELTKNOVSLTCLVKG 357
330 GEYFLAANKM--GSIRHTISVRVKAAPYWLDEPQMLIABEDG-----RLVCRANG- 380
358 YPSDIAVWESNGQENNYKTPPVLDSD--GSFLYSKLTVDKSRMOQGNVFSQVME 415
381 NPKP-SIQMLVNGEP---IEGSPFNPSPREVAGDTIVFRDTQIGSSA-----VYQCMASNE 431

RESULT 90
LAMP HUMAN STANDARD; PRT; 338 AA.
AC Q13449;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Limbic system-associated membrane protein precursor (LSAMP).
OS LSAMP OR LAMP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OK NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96235133; PubMed=8666243;
RA Pimenta A.P., Fischer I., Levitt P.;
RT "cDNA cloning and structural analysis of the human limbic-system-
RL associated membrane protein (LAMP).";
RL Gene 170:189-195(1996).
CC -I- FUNCTION: MEDIATES SELECTIVE NEURONAL GROWTH AND AXON TARGETING.
CC CONTRIBUTES TO THE GUIDANCE OF DEVELOPING AXONS AND REMODELING OF

```

```

CC CC MATURE CIRCUITS IN THE LIMBIC SYSTEM. ESSENTIAL FOR NORMAL GROWTH
CC OF THE HYPOCAMPAL MOSSY FIBER PROJECTION (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -I- TISSUE SPECIFICITY: Expressed on limbic neurons and fiber tracts
CC as well as in single layers of the superior colliculus, spinal
CC chord and cerebellum.
CC -I- SIMILARITY: Belongs to the immunoglobulin superfamily. IGLON
CC family.
CC -I- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC DR EMBL: U41901; AAC50569.1; -.
CC DR PIR: JCA4776; JCA4776.
CC DR Genew: HGNC:6705; LSAMP.
CC DR MIM: 603241; -.
CC DR GO: GO:0007399; P:neurogenesis; TAS.
CC DR InterPro: IPR007110; IG-like.
CC DR InterPro: IPR003598; IG_c2.
CC DR Pfam: PF00047; Ig_3.
CC DR SMART: SM00408; IGC2; 2.
CC DR PROSITE: PS50835; IG_LIKE; 3.
CC KM Immunoglobulin domain; Cell adhesion; glycoprotein; GPI-anchor;
CC Repeat; Signal; Lipoprotein.
CC FT CHAIN 1 28
CC FT SIGNAL 1 28
CC FT PROPEP 316 338
CC FT DOMAIN 29 122
CC FT DOMAIN 132 214
CC FT DOMAIN 219 304
CC FT DISULFID 53 111
CC FT DISULFID 153 197
CC FT DISULFID 239 290
CC FT CARBOHYD 46 40
CC FT CARBOHYD 60 66
CC FT CARBOHYD 136 136
CC FT CARBOHYD 148 148
CC FT CARBOHYD 279 279
CC FT CARBOHYD 287 287
CC FT CARBOHYD 300 300
CC FT CARBOHYD 315 315
CC FT LIPID 315 315
CC SQ SEQUENCE 338 AA; 37308 MW; 0345F286DF5D92F CRC64;

Query Match 5.4%; Score 147; DB 1; Length 338;
Best Local Similarity 22.7%; Pred. No. 0.017;
Matches 75; Conservative 49; Mismatches 129; Indels 76; Gaps 16;

10 LLLVLQALLALPA-----TQGNKVVVGKKQDTVELTCTASOKSIQFHKNSNQIKI 61
14 LVLLRLCLLPGLPVRSVDNFRGDTNITVQSGTALIRVLADKNS-KVALNRSGLIF 72
62 LGNGSFLTKGPKSLNDA--DSRRSLMDQGNFPLINKLIEDSDYICEVDQKE--- 116
73 AGHD-----KMSLDPRVLEKHSI-----EYSLRIQKVNVYBGSYTCVQTOHEKPT 121
117 -EVQLVFG-----LTASDTHLQGSITLTLESPPSSPSVQCR--SPRGKNIQGGKT- 168
122 SQVLLIVGPPKINISSDVIVNGSNVTLVCMANGREPIVTRHRLPTREREGBEY 181
169 LSVSOLELQDSGVTCTVQ-----NOKKVEFKIDYPCPAPEPKSCDKHTTCELLGSP 223
182 LEIGITREOSGKYECKANESVADYQVAVTVNYP--TTEKSKNEAT----- 230
224 SVFLPPPKDTLMSRPEVTCVVVDVSHEDDEVKFNWVDGVEVNA-----KTKPREQ 280

```

Db 231 -----TGROASLKCEASAV-----PAPDFEWMYDDPRINSANGLKSTEGQ 272

QY 281 YNSTYRVSVLTVLHQMVLNGKEYCKKYSNK 311

Db 273 SSLT-----VTNVTEHYGN-----YTCVAANK 295

RESULT 91

NPAS_MOUSE STANDARD; PRT; 1240 AA.

AC Q810U3;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Neurofascin precursor.

GN NPASC.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Swiss Webster; TISSUE=Brain;

RA Dicks P., Montag-Sallaz M., Montag D.;

RT "Expression patterns of Ll-family cell recognition molecules Ll, CHL1, NrcAM, and neurofascin in the mouse brain."

RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: Cell adhesion, ankyrin-binding protein which may be involved in neurite extension, axonal guidance, synaptogenesis, myelination and neuron-glia cell interactions (by similarity).

CC -1- SUBUNIT: Probable constituent of a neurofascin/NRCAM/ankyrin G complex. Associates with the sodium channel beta-1 (SCN1B) and beta-3 (SCN1B) subunits (by similarity).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.

CC -1- SIMILARITY: Contains 6 immunoglobulin type III domains.

CC -1- SIMILARITY: Contains 4 fibronectin type III domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AJ543322; CAD5849.1; -

DR MCD; MG1:2442229; D430023G06R1k.

DR InterPro; IPR003961; FN_III.

DR InterPro; IPR008957; FN_III-like.

DR InterPro; IPR003599; IG_1like.

DR InterPro; IPR007110; IG_1like.

DR InterPro; IPR003598; IG_c2.

DR Pfam; PF00041; fn3; 4.

DR Pfam; PF00047; Ig; 6.

DR SMART; SM00060; FN3; 4.

DR SMART; SM00409; IG; 6.

DR SMART; SM00408; IGC2; 6.

DR PROSITE; PS50835; IG LIKE; 6.

KM Cell adhesion; Repeat; Signal; Transmembrane; Immunoglobulin domain; Glycoprotein.

KM GLYCOPROTEIN.

FT CHAIN 1 24 POTENTIAL.

FT NEUROFASCIN.

FT DOMAIN 25 1110 EXTRACELLULAR (POTENTIAL).

FT TRANSSEM 1131 1131 POTENTIAL.

FT DOMAIN 1132 1240 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 41 137 IG-LIKE C2-TYPE 1.

FT DOMAIN 143 230 IG-LIKE C2-TYPE 2.

FT DOMAIN 244 332 IG-LIKE C2-TYPE 3.

FT DOMAIN 337 424 IG-LIKE C2-TYPE 4.

FT DOMAIN 430 517 IG-LIKE C2-TYPE 5.

FT DOMAIN 521 603 IG-LIKE C2-TYPE 6.

FT DOMAIN 628 720 FIBRONECTIN TYPE-III 1.

FT DOMAIN 727 820 FIBRONECTIN TYPE-III 2.

FT DOMAIN 825 924 FIBRONECTIN TYPE-III 3.

FT DOMAIN 1008 1094 FIBRONECTIN TYPE-III 4.

FT DOMAIN 924 1006 THR-RICH.

FT DISULFID 63 118 POTENTIAL.

FT DISULFID 162 213 POTENTIAL.

FT DISULFID 268 316 POTENTIAL.

FT DISULFID 358 408 POTENTIAL.

FT DISULFID 432 501 POTENTIAL.

FT DISULFID 543 592 POTENTIAL.

FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 778 778 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 866 866 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 881 881 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 1240 AA; 137975 MW; 6DE8935B5B02E965 CRC64;

Query Match 5.4%; Score 147; DB 1; Length 1240;

Best Local Similarity 21.4%; Pred. No. 0.087;

Matches 100; Conservative 66; Mismatches 183; Indels 118; Gaps 26;

QY 6 PFRHLVLVQLALLPA-----TQGNKVVLGKGGTVELCTASQ 45

Db 8 PWHIALILFLSLGGHLEIPMDPSIONELTOPPTTKQSVKHIVDPNNILIECAKG 67

QY 46 KKSIOFHW-KNSNQIKILGNGSFLTK--GPSKLNDRAUSRSIMQGNFPLIKN----- 98

Db 68 NPAPSFMTNNSRPFNIADOPRVSMRRSGTLVIDPFGSGRPREY-EGEYQCFARNKFGT 126

QY 99 -----LKEDSDTYICEVEDQKEVQVLFGTLANSDFHLQGSGLTILSSPPG-SS 150

Db 127 ALSNRIRLQVSKSPLW-----PKENLDPVV-----VOEGAPLTLQCNPPGLPS 170

QY 151 PSVQCSPPRKNIQGGKTLIS-----VSQLELDQSGT-WTCGVLDN-----OKKVEFK 196

Db 171 PVIFWSSSWEPITQDKRVSQGHNGDLVYSNVMLQDQITYSCHARHFHTTLOQKQPT 230

QY 197 IDIVPCPAPRPSKCDTHITCPPELLGSPSVLPFPKPKDTLMISRTPE--VTCVVVDVSH 254

Db 231 LKVLTRGVAERT-----PS-FMYQGSSTSSQMVLRGMDDLIECIASGV--- 273

QY 255 DPEVKENWYDGVVEHNAKTKPREBOYNSYRVSVLTVLHQMVLNGKEYCKKYSNKALP 314

Db 274 -PTPDIAWYKKGDDLPSNKA--FENFNKALRTINVEE-----DSGEYFCIASNKA-- 322

QY 315 APIEKTIS-KAKGQ---REPOVYTLPPSRDELTKQVSLTCLVKGYPSDIAVEMESNG 370

Db 323 GSIRHTISRVKAPYWLDEPKVLLAPGEDG-----RLVCRANG-NRP-P-TVQMMVNG 374

QY 371 QPENNYKTPPVLDSD--GSFLLYSKLTVDKSRMOQGNVFCSVME 415

Db 375 EF---LQSAPEPNPREVAQDTIIFRDTQISSRA-----VYQCNTSNE 413

RESULT 92

NPEN_MOUSE STANDARD; PRT; 1242 AA.

AC Q9QZS7;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Nephxin precursor (Renal glomerulus-specific cell adhesion receptor).

GN NPXS1 OR NPEN.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RN [1]


```

FT CAROHVD 1029 1029 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CAROHVD 1072 1072 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CAROHVD 1106 1106 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 1179 1182 Missing (in isoform 2).
FT VARSPLIC 1179 1182 Missing (in isoform 2).
SQ SEQUENCE 1259 AA; 140934 MW; 0F12A7C4415F3C08 CRC64;

Query Match 5.4%; Score 146; DB 1; Length 1259;
Match Local Similarity 24.4%; Pzed. No. 0.1;
Match 99; Conservative 43; Mismatches 138; Indels 126; Gaps 24

OY 10 LLLVQLALLPAAATGKNNKVLGKKDPYELTCTASOKSIO--FMHKSNNQKILGNQGS 67
DB 507 ILANIQVNAETITGSPSTIKKARVYTFQASFDPSLQASITWRGDR----- 557
OY 68 FLTGPSKLNDRAADRSRLMDQGNFPLIKMLKIEDSDTYIC---EVEDQKEVOLLIVF 123
DB 558 -----DLQERGDSDKVFIEDQ--LVTKSLDYSDDQDYGCVASTELDEVESRAQLLV 608
OY 124 GLTAN-----SDHLQGGSLTLTLESPGSSPSVQCRSP-----RCNNIGGKTL 169
DB 609 GSPGFVPHLELSDRHLKLKQSQVHLWS-----SPADHNNSPIEKYDIEBEDKEMAPKWF 662
OY 170 SVSOLELDDSGWTCVTLQNGKVE--FKDIVPCPAP--EPSCOKHTCPRELLGGPSVF 226
DB 663 SLGKR-----PGNQSTTLKLSPIYHYTFKVAINKRGPEPSVASETVVTP----- 710
OY 227 LFPYPRKDTLMISRTPEVTVCVVVDVSHEDPEVK--FNWYVDGVGVHNAKTP--REEQYN 282
DB 721 -----AAPEKN--PVDVRGEONEENMMNVITM-----KPLRMDMN 743
OY 283 S---TYRVSVTLVTLHQMVLN--GKEYKCK-----VSNKALPAPIEKTISTPAKQP 328
DB 744 APQIQYRV-----QMRPLGKQETWKEQTSVDPFLVASTSTFVFPEIKVQAVNNQG 794
OY 329 R--EPOV---YN-----LPPRSBLTKNQVSLTGLVNGFYSDIA 363
DB 795 KGPEPOVTITGYSGEDYPOVSPLEEDITIFN--SSTLVLR--WRPVDA 838

RESULT 95
LAC_CHICK STANDARD; PRT; 103 AA.
ID_LAC_CHICK STANDARD; PRT; 103 AA.
AC P20763;
DT 01-FEB-1991 (rel. 17, Created)
DT 01-FEB-1991 (rel. 17, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Ig lambda chain C region.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxId=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87218480; PubMed=3107981;
RA Parvati R., Ziv E., Lennner F., Tel-Or S., Burestein Y., Schechter I.;
RT "Analyses of chicken immunoglobulin light chain cDNA clones indicate
a few germ-line V lambda genes and allotypes of the C lambda locus.",
EMBO J. 6:97-102(1987).
RL
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL, X04768; CAA28461.1; -
DR PIR, B2167; B26167.
DR HSP, P01842; 7FAB.
```

DR	InterPro; IPR007110; Ig-1-like.	
DR	InterPro; IPR003597; Ig_c1.	
DR	InterPro; IPR003006; Ig_MHC.	
DR	Pfam; PF00047; Ig; 1.	
DR	SMART; SMO0407; Igc1; 1.	
DR	PROSITE; PS50835; Ig_LIKE; 1.	
DR	PROSITE; PS00290; Ig_MHC; 1.	
KM	Immunoglobulin domain; Immunoglobulin C region.	
FT	NON_TER	1
FT	DOMAIN	1
FT	DISULFID	6 99
FT	DISULFID	28 85
FT	DISULFID	103 103
FT	VARIANT	90 90
SO	SEQUENCE	103 AA; 11361 MW; 77BFB341B511B91B2 CRC64;
Query Match	5.4%; Score 145.5; DB 1; Length 103;	
Beat Local Similarity	32.2%; Pred. No. 0.0048;	
Matches	29; Conservative 19; Mismatches 39; Indels 3; Gaps 3	
OY	337 QPR-EPQVYTLPPSRDELTK-NQVSLTGLVKGFPYSIDIAVEWESNGQPENNYKTTTPPID 384	
DB	1 QPKVAPRTLLPPPSKEELNEATKATLVCLINDFYSPVTVMDVIGSTRSG-ETTAPORO 59	
OY	385 SDGSFFLYSKLTVDKSRWQCGNVFSCSYMH 414	
DB	60 SNSQYMASSYLSLSASDWSHETVYCRVTH 89	
RESULT 96		
KACB_RAT	STANDARD; PRT; 106 AA.	
ID_KACB_RAT	P01835;	
AC	21-JUL-1986 (Rel. 01, Created)	
DT	21-JUL-1986 (Rel. 01, Last sequence update)	
DT	10-OCT-2003 (Rel. 42, Last annotation update)	
DE	Ig kappa chain C region, B allele.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxId=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=Louvain;	
RX	MEDLINE=82082587; PubMed=6273908;	
RA	Sheppard H.W., Gutman G.A.;	
RT	"Allelic forms of rat kappa chain genes: evidence for strong	
RT	selection at the level of nucleotide sequence.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 78:7064-7068(1981).	
RN	[2]	
RP	SEQUENCE (BENCE-JONES PROTEIN S211).	
RX	MEDLINE=75212238; PubMed=8076530;	
RA	Starace V., Querinjean P.;	
RT	"The primary structure of a rat kappa Bence Jones protein:	
RT	phylogenetic relationships of V- and C-region genes.";	
RL	J. Immunol. 115:59-62(1975).	
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like domain.	
DR	PIR; A93901; KIRMB.	
DR	HSSP; P01842; 2MCG.	
DR	InterPro; IPR007110; Ig-1-like.	
DR	InterPro; IPR003597; Ig_c1.	
DR	InterPro; IPR003006; Ig_MHC.	
DR	Pfam; PF00047; Ig; 1.	
DR	SMART; SMO0407; Igc1; 1.	
DR	PROSITE; PS50835; Ig_LIKE; 1.	
DR	PROSITE; PS00290; Ig_MHC; 1.	
KM	Immunoglobulin domain; Immunoglobulin C region.	
FT	NON_TER	1
FT	DOMAIN	1
FT	DISULFID	5 102
FT	DISULFID	106 86
FT	DISULFID	106 106
FT	CONFLICT	2 2
FT	CONFLICT	30 30
FT	CONFLICT	48 48
INTERCHAIN (WITH A HEAVY CHAIN).		
D -> N (IN REF. 2).		
N -> K (IN REF. 2).		
MISSING (IN REF. 2).		

```

FT CONFLICT 79 79 E -> Q (IN REF. 2).
FT CONFLICT 87 87 E -> Q (IN REF. 2).
FT CONFLICT 98 98 V -> VM (IN REF. 2).
FT CONFLICT 100 100 S -> N (IN REF. 2).
SQ SEQUENCE 106 AA: 11601 MW: 4CFA7CA820D1CA36 CRC64;

Query Match
Best Local Similarity 31.5%; Pred. No. 0.0053;
Matches 29; Conservative 22; Mismatches 29; Indels 12; Gaps 2;

Qy 331 PQTTLPSRDELTONQSLCTVKGFPSPDIAMWENSGPENNKTTPVLD----- 385
Db 5 PTVSIFPSTQLATGASVCLMNNFYPRDISVTKMIDGERRD-----GVLDSVTDQD 59
Qy 386 --DGSFPLYSKLTVDKSRMOQGNVPSGSMHE 415
Db 60 SKDSTYSMSSTLSLTAKDYESHNLTYTCVYHK 91

RESULT 97
AMAL_DROME STANDARD: PRT: 333 AA.
AC PLS364; Q9V3A5;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Amalgam protein precursor.
GN AMA OR BG:DS00276.6 OR CG2198.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=89028670; PubMed=3141062;
RA Seeger M.A., Haffley L., Kaufman T.C.;
RT "Characterization of amalgam: a member of the immunoglobulin
RT superfamily from Drosophila.";
RL Cell 55:589-600(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Celniker S.E., Pfeiffer B.D., Knafels J., Martin C.H., Mayeda C.A.,
RA Palazzolo M.J.;
RT "Complete sequence of the Antennapedia complex of Drosophila.";
RT submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M.P., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Buttle K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Charry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Houchin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

```

RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobary C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Puri V., Reese M.G.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spadling A.C., Stapleton M., Strong R., Sun E.,
RA Svrcek R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Waberman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guerin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (potential).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M23561; AAA28367.1; -.
DR EMBL: AE001572; AAD19797.1; -.
DR EMBL: AE003674; AAF54084.1; -.
DR EMBL: AY051911; AAK93335.1; -.
DR PIR: A31923; A31923.
DR FLYBase: FBgn0000071; Ama.
DR GO: GO:0005886; C:Plasma membrane; IDA.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003598; IG_c2.
DR Pfam: PF00047; IG_3.
DR SMART: SMO0408; IGC2; 2.
DR PROSITE: PSS00835; IG_LIKE; 3.
KW Immunoglobulin domain; Glycoprotein; Membrane; GPI-anchor; Signal;
KW Repeat.
FT 1 23 POTENTIAL.
FT CHAIN 24 ? AMALGAM PROTEIN.
FT PROPEP 2 333 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 25 128 IG-LIKE V-TYPE.
FT DOMAIN 139 223 IG-LIKE C2-TYPE 1.
FT DOMAIN 230 323 IG-LIKE C2-TYPE 2.
FT DISULFID 46 117 PROBABLE.
FT DISULFID 161 208 PROBABLE.
FT DISULFID 251 307 PROBABLE.
FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 308 308 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 333 AA: 36387 MW: F644753DE3DB25P1 CRC64;

Query Match
Best Local Similarity 19.9%; Pred. No. 0.024;
Matches 70; Conservative 57; Mismatches 127; Indels 97; Gaps 14;

```

Qy 10 LLLVLQAL-----LPAAQGNKRVVLGKGDVETLCTASQKSIQPHW-----KNSNQ 58

```

Db      10 LIFCLALSLDSVLSAPVLSISKDVASVSDSVFPCVTEVEVQGLSVNAKRSRSEDTS 69
Qy      59 IKI-----LGNQ--GSFLTKPSPKLNDRADSRSLMDQGNPLLIKNLKDSDPTIYC 109
Db      70 VVLISMNRLISLPQORVNVVTEGPRT-----GSAITPFIQNIENVSDMGPRYEC 117
Qy      110 EV-----EDQKEVQLL--VFGLTNSDTHLQOQS/TLTLESPGSSPSV----- 153
Db      118 QVLVSATKTKLKLSDIKTPPIAIENTPKSTLVTEGQNLCTCHANGPKFTISWARRH 177
Qy      154 QCRSPKNGIKQSGKTLVSQLELQDSGTWTCTVLON-----OKKVEFKIDIVPCP 203
Db      178 NAWMPAGGHLLAPPTLRIRSVHMDGYYC-IAQNGEGQPKRLLRVEVEFRPQLA--- 233
Qy      204 APEPKSCDKHTHTPELLGGSVLFPPEPKQTMISTRTPEVTVVVDVSHEDPEVKFNMY 263
Db      234 -----VORPKIAGMWSHSLSECVQGY-----PAPTVVWH 264
Qy      264 VDGVEVNAKTKPREQVNSTYRVSVL---TVLHQMNLGKREYKCKVSNK 311
Db      265 KNGVPLQSRHHEVAVNTASSGTTTSVLRIQSVGEEDF---GDYICNAIK 312

RESULT 98
VCAM_MOUSE STANDARD: PRT: 739 AA.
ID VCAM_MOUSE
AC P29533;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vascular cell adhesion protein 1 precursor (V-CAM 1).
GN VCAM1 OR VCAM-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=FVB; TISSUE=Lung;
RX MEDLINE=92161437; Pubmed=1371918;
RA Hession C., Moy P., Tizard R., Chisholm P., Williams C., Wysk M.,
RA Burkly L., Miyake K., Kincaide P., Lobb R.;
RT "Cloning of murine and rat vascular cell adhesion molecule-1.";
RL Biochem. Biophys. Res. Commun. 183:163-169(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lymph node;
RX MEDLINE=93246254; Pubmed=7683304;
RA Arai M., Arai K., Vassalli P.;
RT "Cloning and sequencing of mouse VCAM-1 cDNA.";
RL Gene 126:261-264(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=129; TISSUE=Embryo;
RX MEDLINE=94117008; Pubmed=7507076;
RA Cybulsky M.I., Allen-Mocamed M., Collins T.;
RT "Structure of the murine VCAM1 gene.";
RL Genomics 18:387-391(1993).
RN [4]
RP SEQUENCE OF 1-693 FROM N.A. (ISOFORM 1).
RC STRAIN=N1H Swiss, and 129/Sv;
RA Kumar A.G., Dai Y.X., Kozak C.A., Mime M.P., Gotto A.M. Jr.,
RA Ballantyne C.M.;
RN Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=FVB; TISSUE=Lung;
RX MEDLINE=9332042; Pubmed=7682556;
RA Moy P., Lobb R., Tizard R., Olson D., Hession C.;
RT "Cloning of an inflammation-specific phosphatidyl inositol-linked
RT form of murine vascular cell adhesion molecule-1.";
RL J. Biol. Chem. 268:8835-8841(1993).

```

```

RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=95015899; Pubmed=7523515;
RA Kumar A.G., Dai Y.X., Kozak C.A., Mime M.P., Gotto A.M.,
RA Ballantyne C.M.;
RT "Murine VCAM-1. Molecular cloning, mapping, and analysis of a
RT truncated form.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4088-4098(1994).
RN [7]
RP SEQUENCE OF 311-345 FROM N.A. (ISOFORM 2).
RC STRAIN=FVB/N; TISSUE=Kidney;
RX MEDLINE=93317595; Pubmed=7687058;
RA Terry R.W., Kwee L., Levine J.F., Labow M.A.;
RT "Cyclokin induction of an alternatively spliced murine vascular cell
RT adhesion molecule (VCAM) mRNA encoding a
RT glycosylphosphatidylinositol-anchored VCAM protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:5919-5923(1993).
RN [8]
RP SEQUENCE OF 1-21 FROM N.A.
RC TISSUE=Endothelial cells;
RA Korenaga R., Ando J., Teubol H., Kamiya A.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION
CC IN LEUCOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1
CC INTEGRIN VIA4 ON LEUCOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL
CC TRANSDUCTION. THE VCAM1/VLA4 INTERACTION MAY PLAY A
CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUCOCYTE
CC EMIGRATION TO SITES OF INFLAMMATION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (LONG ISOFORM) OR
CC ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (SHORT ISOFORM).
CC -1- ALTERNATIVE PRODUCTS:
CC Name=1; Synonyms=Long;
CC Name=2; Synonyms=Short;
CC IsoId=P29533-1; Sequence=Displayed;
CC IsoId=P29533-2; Sequence=VSP 002581, VSP 002582;
CC Name=2; Synonyms=Short;
CC -1- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as
CC well as on macrophage-like and dendritic cell types in both normal
CC and inflamed tissue.
CC -1- PTM: The GPI-anchor is located on position 319 of isoform 2.
CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M8487; AAA40545.1; -
CC EMBL; X67783; CAA47989.1; -
CC EMBL; L22355; AAA16921.1; -
CC EMBL; L22301; AAA16921.1; JOINED.
CC EMBL; L22349; AAA16921.1; JOINED.
CC EMBL; L22350; AAA16921.1; JOINED.
CC EMBL; L22351; AAA16921.1; JOINED.
CC EMBL; L22352; AAA16921.1; JOINED.
CC EMBL; L22353; AAA16921.1; JOINED.
CC EMBL; L22354; AAA16921.1; JOINED.
CC EMBL; L22355; AAA16921.1; JOINED.
CC EMBL; L22350; AAA16920.1; -
CC EMBL; L22301; AAA16920.1; JOINED.
CC EMBL; L22349; AAA16920.1; JOINED.
CC EMBL; U12878; AAB60659.1; ALT_SEQ.
CC EMBL; U12879; AAB60660.1; ALT_SEQ.
CC EMBL; U12880; AAB60661.1; ALT_SEQ.
CC EMBL; U12874; AAB60662.1; ALT_SEQ.
CC EMBL; U12871; AAB60663.1; ALT_SEQ.
CC EMBL; U12883; AAB60664.1; ALT_SEQ.
CC EMBL; U12881; AAA80010.1; ALT_SEQ.
CC EMBL; U12882; AAA80011.1; ALT_SEQ.
CC EMBL; U12875; AAA80012.1; ALT_SEQ.

```


FT	CHAIN	29	315	LIMBIC SYSTEM-ASSOCIATED MEMBRANE PROTEIN
FT	PROPEP	316	338	REMOVED IN MATRE FORM (POTENTIAL).
FT	DOMAIN	29	122	IG-LIKE C2-TYPE 1.
FT	DOMAIN	132	214	IG-LIKE C2-TYPE 2.
FT	DOMAIN	219	304	IG-LIKE C2-TYPE 3.
FT	DISULFID	53	111	POTENTIAL.
FT	DISULFID	153	197	POTENTIAL.
FT	DISULFID	239	290	POTENTIAL.
FT	CARBOHYD	40	40	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	66	66	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	148	148	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	279	279	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	287	287	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	300	300	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	315	315	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	LIPID	315	315	GPI-anchor amidated asparagine (Potential).
SO	SEQUENCE	338 AA;	37324 MW;	0B76AFDDB6A39BB6 CRC64;
Query Match		5.3%;	Score 144;	DB 1; Length 338;
Best Local Similarity		22.4%;	Pred. No. 0.026;	
Matches		74;	Conservative	50; Mismatches 129; Indels 78; Gaps 16.
QY	10	LLVLVQLALLPA-----TQGNKVVLGKKGDVLELTCTASQKKSIOFHKNSNQIKI	61	
DB	14	LVLRLCLLPGLTPVRSDVDFNRGTDNITVQGDPAILRCVDEDKNS-KVAMLNRSGLIF	72	
QY	62	LGNQGSFLTKGSPSKINDRA--DSRSIWDQGNFPLIINKIKEDSDTYICEVEDKE---	116	
DB	73	AGHD-----KMSLDIPVELEKRIAL---EYSLRIKVDVYDEGSYTCVSOTHEPKT	121	
QY	117	-EVOLVVG-----ITANSDFHLQGSITLTLESPPGSSPSVQCR--SPRGKNIQGKT-	168	
DB	122	SOVLIVIVQPKRISISSDYVNBESNTVLVMANGREPEVITWHLPLPLREFEGEEHY	181	
QY	169	LSVSQLEIQDSGTWTCTVLQ-----NQKKVEFKIDIVCPAPRPSCDKHTCPELLGAP	223	
DB	182	LEIIGITREQSGKYECKAANEVSADVQKVATVYYP-TITESKSNBAT-----	230	
QY	224	SVFLFPPKPKOTLMISRTPEYTCVVVDVSHPEPKFMVYDVGVVHNA--KTYPREEQ	280	
DB	231	-----TGRQASLKCEASAV-----PADDFEWMYRDDTRINSANGLEIKSTEGQ	272	
QY	281	YNSTRVYVSVLTVLHODWLNGKEYCKYSNK	311	
DB	273	SSLT-----VNVTEHYGN---YTCVAANK	295	
RESULT	100			
ID	NFAS_RAT	STANDARD;	PRT:	1240 AA.
AC	P97685; P97684; Q91260;			
DT	15-MAR-2004 (Rel. 43, Created)			
DT	15-MAR-2004 (Rel. 43, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Neurofascin precursor.			
GN	NPASC.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_Taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 3).			
RC	STRAIN=leucaris;			
RA	Tat S., Collinson J.M., Brophy P.J.;			
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE OF 25-1240 FROM N.A. (ISOFORMS 1; 2 AND 3), AND SUBUNIT.			
RC	TISUE=Brain;			
EX	MEDLINE=97103184; PubMed=8947556;			
RA	Davis J.Q., Lambert S., Bennett V.;			

RT "Molecular composition of the node of Ranvier: identification of
RT ankyrin-binding cell adhesion molecules neurofascin (mucin+/chirid
RT FNIII domain-) and NrCAM at nodal axon segments.";
RL J. Cell Biol. 135:1355-1367(1996).
[3]
RP FUNCTION OF ISOFORM 2/3.
RX MEDLINE=98220650; Pubmed=9562181;
RA Collinson J.M., Marshall D., Gillespie C.S., Brophy P.J.;
RT "Transient expression of neurofascin by oligodendrocytes at the onset
RT of myelinogenesis: implications for mechanisms of axon-glia
RT interaction.";
RL Glia 23:11-23(1998).
[4]
RP FUNCTION OF ISOFORMS 1 AND 2/3, INDIRECT ASSOCIATION WITH CNTNAP1,
RP PHOSPHORYLATION, AND TISSUE SPECIFICITY.
RX MEDLINE=20391985; Pubmed=10931875;
RA Tait S., Gunn-Moore F., Collinson J.M., Huang J., Lubetcki C.,
RA Pedraza L., Sherman D.L., Colman D.R., Brophy P.J.;
RT "An oligodendrocyte cell adhesion molecule at the site of assembly of
RT the paranodal axo-glia junction.";
RL J. Cell Biol. 150:657-666(2000).
[5]
RP INTERACTION WITH SODIUM CHANNEL BETA-1 AND BETA-3 SUBUNIT.
RX MEDLINE=21365577; Pubmed=11470822;
RA Racciliffe C.F., Westendorp R.E., Curtis R., Caterall W.A.;
RT "Sodium channel beta1 and beta3 subunits associate with neurofascin
RT through their extracellular immunoglobulin-like domain.";
RL J. Cell Biol. 154:427-434(2001).
CC -1- FUNCTION: Cell adhesion, ankyrin-binding protein which may be
CC involved in neurite extension, axonal guidance, synaptogenesis,
CC myelination and neuron-glia cell interactions. Isoforms 2/3 may
CC be responsible for mediating and signaling axon-glia interaction
CC during the early stages of myelination.
CC -1- SUBUNIT: Probable constituent of a neurofascin/NRCAM/ankyrin G
CC complex. Associates with the sodium channel beta-1 (SCNB1) and
CC beta-3 (SCNB3) subunits. Associates to beta-1 subunit in
CC developing axons as early as postnatal day 5, during the period
CC that nodes of Ranvier are forming. Isoform 2/3 is likely to
CC interact with axonal proteins in close association with CNTNAP1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 1
CC colocalizes with ankyrin G at the nodes of Ranvier. Isoform 2/3 is
CC a glial component of the paranodal axo-glia junction.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1; Synonyms=NP185, 186 kDa isoform;
CC IsoId=P97685-1; Sequence=Dsplayed;
CC Name=2; Synonyms=NP155, 155 kDa isoform;
CC IsoId=P97685-2; Sequence=VSP_050416, VSP_050417, VSP_050418,
CC VSP_050419;
CC Name=3; Synonyms=NP155, 155 kDa isoform;
CC IsoId=P97685-3; Sequence=VSP_050416, VSP_050417, VSP_008941,
CC VSP_050418, VSP_050419;
CC -1- TISSUE SPECIFICITY: Isoform 1 is expressed at Nodes of Ranvier
CC while isoform 2/3 is expressed in unmyelinated axons.
CC -1- DEVELOPMENTAL STAGE: Strongly but transiently up-regulated in
CC oligodendrocytes at the onset of myelinogenesis. Once these last
CC have engaged their target axons, expression declines
CC precipitously.
CC -1- PFM: Isoform 2/3 is phosphorylated at P12. Dephosphorylation is
CC required for ankyrin binding.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC Li/neurofascin/NGCAM family.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 4 fibronectin type III domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AY061639; AAL27854.1; -.
 DR EMBL: U81035; AAB47753.1; -.
 DR EMBL: U81036; AAB47754.1; -.
 DR HSPSP; P20241; 1CPE.
 DR InterPro; IPR003961; FN_III-like.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR007110; IG_1like.
 DR InterPro; IPR003598; IG_C2.
 DR Pfam; PRO0047; fn3; 4.
 DR Pfam; PRO0047; fn3; 6.
 DR SMART; SM00060; FN3; 3.
 DR SMART; SM00408; IG_C2; 4.
 DR PROSITE; PSS00835; IG_LIKE; 6.
 DR Cell adhesion; Repeat; Signal; Transmembrane; Immunoglobulin domain;
 KM Glycoprotein; Alternative splicing; Phosphorylation.
 FT SIGNAL 1 24
 FT CHAIN 25 1240
 FT DOMAIN 25 1110
 FT TRANSMEM 1111 1131
 FT DOMAIN 1132 1240
 FT DOMAIN 41 137
 FT DOMAIN 143 230
 FT DOMAIN 244 332
 FT DOMAIN 337 424
 FT DOMAIN 429 517
 FT DOMAIN 521 603
 FT DOMAIN 628 720
 FT DOMAIN 727 830
 FT DOMAIN 825 918
 FT DOMAIN 1008 1094
 FT DOMAIN 913 1006
 FT DISULFID 63 118
 FT DISULFID 162 213
 FT DISULFID 268 316
 FT DISULFID 358 408
 FT DISULFID 452 501
 FT DISULFID 543 592
 FT CARBOHYD 305 305
 FT CARBOHYD 409 409
 FT CARBOHYD 446 446
 FT CARBOHYD 483 483
 FT CARBOHYD 752 752
 FT CARBOHYD 778 778
 FT CARBOHYD 866 866
 FT CARBOHYD 881 881
 FT CARBOHYD 1015 1015
 FT CARBOHYD 1026 1026
 FT CARBOHYD 1047 1047
 FT CARBOHYD 1101 1101
 FT VARSPLIC 31 36
 FT VARSPLIC 236 236
 FT VARSPLIC 611 625
 FT VARSPLIC 824 824
 FT VARSPLIC 928 1096
 FT CONFLICT 482 482
 FT CONFLICT 675 675
 FT CONFLICT 699 699
 FT CONFLICT 763 763
 FT SEQUENCE 1240 AA; 138003 MW; 636A187BC3772513 CRC64;
 Query Match 5.3%; Score 143.5; DB 1; Length 1240;
 Best Local Similarity 21.4%; Pred. No. 0.15;
 Matches 106; Conservative 67; Mismatches 193; Indels 129; Gaps 28;

QY 6 PFRHLVLVLQALDLPAA-----TQGNKVVIGKKGGDTVELTCTASQ 45
 DB 8 FVHVVALILFLSLGGAIEIPMDSIQNELTQPTITKQSVKDIIVDRDNILIECEAKG 67
 QY 46 KKSIOFHW-KKSNQIKILGNQGSFLTK--GSKLNDPADSRSLMDQGNFLLIKN----- 98
 DB 68 NPASFHWTRNSRFNFNAKQPRVSMRRRSGTLVIDFSGGRPEBY-EGEYOCFARNKFGT 126
 QY 99 -----LKIDSDTYICEVEDQEEVQLVFGLTANSDTLLQGGSLTLTLESPPG-SS 150
 DB 127 ALSNRIRLQVSKSLW-----PKENIDPV-----VQGAFLTLQCNPPGLPS 170
 QY 151 PSVQCRSPRGKNIQGGKTLT-----VSQLEIQDSGT-WTCTVLQN-----QKVEFK 196
 DB 171 FVIFWMSMSEPIIQDKRVSQGHNGDLVFSNVMQLDMQTDVSCNARFHTHTIQKXNFT 230
 QY 197 IDIYPCAPPEPKSCDKHTTCPELLGGSVFLFPKPKDITLMISTPE--VTCVVVDVSH 254
 DB 231 LKVLITTRGVAERT-----PS-FWYPGTSSSQVVLKGMDLLECLASGV--- 273
 QY 255 DPEYKFNWYVDGVVHNAKTKPREOYNSYTRVSVLTVLHQDWLNGKEYCKVSNKALP 314
 DB 274 -PTEDIANVKKKGDLPSDKAK--FENFKALRITNVSE-----DSGEYFCLASNGM-- 322
 QY 315 APIEKTIS-KAKGP--REPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVESWNG 370
 DB 323 GSIRHTISVRKAAPYMLDEPKNLLIAPGEDG-----RLVCRANG-NPKP-TVQMTLVNG 374
 QY 371 QPENNYKTPPVLDSD--GSFPLYSKLTVDKSRMOQGVFSCSVNHEALNHHYQKSL-- 426
 DB 375 DP--LQAPNPVREVAAGDTIIFRDQIISRA-----VYQCNYSNE--HGVLNANAFVS 424
 QY 427 -----SLSPGLQL 434
 DB 425 VLVDVPPRLSPRNQL 439

Search completed: August 3, 2004, 13:09:30
 Job time : 13.0314 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 13:00:24 ; Search time 11.7668 Seconds
(without alignments)
4185.504 Million cell updates/sec

Title: SEQ8
Perfect score: 2702
Sequence: 1 NMRGVPRHLLVLQALLP.....VISFLGLGLVAVCIARTR 512

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283166 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283166

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 125 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1227.5	45.4	330	1 GHNU	Ig gamma-1 chain C
2	1221	45.2	374	2 S69339	Ig heavy chain V r
3	1203.5	44.5	255	4 S31866	Ig gamma-1 chain C
4	1173.5	43.4	234	2 PTO207	Ig gamma chain C r
5	1157	42.8	377	2 A23511	Ig gamma-3 chain C
6	1155	42.7	377	2 A60764	Ig gamma-3 chain C
7	1153	41.9	289	1 G3HWT	Ig gamma-3 heavy C
8	1130.5	41.8	326	1 G2HU	Ig gamma-2 chain C
9	1119	41.4	327	1 G4HU	Ig gamma-4 chain C
10	1023	37.9	458	1 RMHT4	T-cell surface gly
11	955.5	35.4	398	1 G3MSM	Ig gamma-3 chain C
12	932.5	34.5	393	1 G1MSM	Ig gamma-1 chain C
13	915.5	33.9	470	1 S22080	Ig heavy chain pre
14	906.5	33.5	399	1 G2MSAM	Ig gamma-2a chain
15	901	33.3	328	2 I47159	Ig gamma-2a chain
16	899	33.3	323	1 GHRB	Ig gamma chain C r
17	898	33.2	327	2 I47162	Ig gamma 4 chain c
18	898	33.2	328	2 I47160	Ig gamma 2b chain
19	894.5	33.1	329	1 G2GP	Ig gamma-2 chain C
20	879.5	32.5	405	1 I47158	Ig gamma-1 chain C
21	877.5	32.4	432	1 G2MSBM	Ig gamma-2b chain
22	875	32.4	432	1 RMCZT4	T-cell surface gly
23	873	32.3	329	2 I47161	Ig gamma 3 chain c
24	857.5	31.7	329	1 G3MSC	Ig gamma-3 chain C
25	854.5	31.6	444	2 PC4436	monoclonal antibod
26	854.5	31.6	472	2 S31459	Ig gamma-1 chain -
27	850.5	31.5	469	2 S37483	Ig gamma-2a chain
28	838	31.0	308	2 C30554	Ig heavy chain C r
29	833	30.8	333	2 PS0018	Ig gamma-2b chain

30	825	30.5	446	2 S40295	Ig gamma-2a chain
31	823.5	30.5	326	2 PS0017	Ig gamma-1 chain C
32	823	30.5	329	2 S00847	Ig gamma-2c chain
33	818.5	30.3	324	1 G1MS	Ig gamma-1 chain C
34	818	30.3	330	1 G2MS11	Ig gamma-2b chain
35	805.5	29.8	330	1 G2MSA	Ig gamma-2a chain
36	801	29.6	335	1 G2MSAB	Ig gamma-2a chain
37	793.5	29.4	475	2 S01321	Ig gamma-2b chain
38	792.5	29.3	322	2 PS0019	Ig gamma-2a chain
39	790	29.2	432	1 RMWQ74	T-cell surface gly
40	774	28.6	327	2 S06611	Ig gamma-2 chain C
41	702	26.0	180	2 I46732	Ig gamma heavy cha
42	601.5	22.3	459	2 A46254	CD4 precursor - ra
43	581.5	21.5	432	2 S30193	T-cell surface gly
44	574.5	21.3	218	2 A36040	Ig heavy chain V-I
45	572.5	21.2	249	2 S69340	Ig heavy chain VHI
46	566	20.9	152	2 S14236	Ig gamma-1 chain C
47	495	18.3	457	2 A27449	T-cell surface gly
48	475	17.6	457	1 RMWST4	T-cell surface gly
49	411.5	15.2	572	2 B46529	Ig Y heavy chain (
50	402.5	14.9	549	2 S04845	Ig heavy chain pre
51	389	14.4	627	2 S14683	Ig mu chain precu
52	388	14.4	388	1 EHMS	Ig epsilon chain C
53	386.5	14.3	474	2 S15590	Ig heavy chain - h
54	385	14.2	453	2 S37768	Ig mu chain C regi
55	384	14.2	592	2 S25705	Ig mu chain - shee
56	381	14.1	473	1 MHTDM	Ig mu chain C regi
57	381	14.1	548	2 S38864	Ig epsilon chain C
58	376.5	13.9	452	1 MHRU	Ig mu chain C regi
59	374	13.8	448	2 S03186	Ig heavy chain C r
60	373.5	13.8	476	1 MHMSM	Ig mu chain C regi
61	370.5	13.7	429	1 EHRT	Ig epsilon chain C
62	369.5	13.7	479	1 MHRBM	Ig mu chain C regi
63	369	13.7	450	1 HDG	Ig mu chain C regi
64	365.5	13.5	391	1 MHRUBT	Ig mu heavy chain
65	363.5	13.5	423	1 EHMS	Ig epsilon chain C
66	363	13.4	455	2 A24976	Ig mu chain C regi
67	362	13.4	455	1 MHMS	Ig mu chain C regi
68	361.5	13.4	426	2 I36948	Ig epsilon-chain
69	361.5	13.4	428	1 EHRU	Ig epsilon chain C
70	361	13.4	504	2 S00390	Ig gamma chain (C1
71	360	13.3	458	1 MHRB	Ig mu chain C regi
72	357	13.2	71	2 I60082	CD4 receptor - hum
73	354.5	13.1	457	2 S03961	Ig mu chain C regi
74	353	13.1	343	2 S25644	Ig mu chain C regi
75	349	12.9	453	1 MHRH	Ig mu chain C regi
76	345.5	12.8	453	2 C31933	Ig mu chain C regi
77	331.5	12.3	112	2 B30503	Ig gamma-2a chain
78	325.5	12.0	577	2 I50731	Ig heavy chain - n
79	321.5	11.9	433	2 S31436	Ig epsilon chain -
80	316.5	11.7	240	2 A39016	Ig mu chain C regi
81	312	11.5	438	1 HVRK2	Ig mu chain - lepi
82	308.5	11.4	474	2 I50830	T-cell surface gly
83	305.5	11.3	99	2 S21461	Ig mu chain C regi
84	301.5	11.2	461	1 HVRKCO	Ig alpha chain - c
85	301.5	11.2	585	2 A46507	Ig alpha chain - c
86	301.5	11.2	684	2 S60266	novel antigen rece
87	300	11.1	342	2 I47175	Ig alpha chain C r
88	297	11.0	339	2 S09264	Ig alpha chain C r
89	296	11.0	357	2 S09269	Ig alpha chain C r
90	296	11.0	367	1 MHCH	Ig mu chain C regi
91	296	11.0	454	2 A46532	Ig mu chain C regi
92	295.5	10.9	340	2 S09272	Ig alpha chain C r
93	294.5	10.9	343	2 B22360	Ig alpha-2 chain C
94	294	10.9	438	1 HVRKCS	Ig mu chain C regi
95	293	10.8	299	1 AHRB	Ig alpha chain C r
96	292	10.8	352	2 S09276	Ig alpha chain C r
97	292	10.8	352	2 S09266	Ig alpha chain C r
98	290	10.7	568	2 A34891	Ig heavy chain pre
99	289	10.7	357	2 S09265	Ig alpha chain C r
100	286.5	10.6	340	2 I56230	Ig alpha-2 chain -
101	286.5	10.6	370	1 HVRKCS	Ig mu chain C regi
102	285.5	10.6	348	2 S09270	Ig alpha chain C r

```

103      284      10.5      244      2      S12328      Ig heavy chain C r
104      282.5     10.5      348      2      S09273      Ig alpha chain C r
105      282      10.4      347      2      S09274      Ig alpha chain C r
106      281      10.4      393      1      HVRKCI      Ig mu chain C regi
107      280.5     10.4      99      2      S21462      T-cell surface gly
108      279      10.3      352      2      S05500      Ig alpha-1 chain C
109      277.5     10.3      353      1      ALHU       Ig alpha-1 chain C
110      277      10.3      340      1      A2HU       Ig alpha-2 chain C
111      276.5     10.2      568      2      A45804      Ig mu chain C regi
112      275.5     10.2      88      2      A35003      Ig gamma-2b chain
113      275      10.2      357      2      S09267      Ig alpha chain C r
114      268.5     9.9      360      2      S09271      Ig alpha chain C r
115      264.5     9.8      344      1      AHW5       Ig alpha chain C r
116      263.5     9.8      358      2      S09268      Ig alpha chain C r
117      258      9.5      580      2      A46538      Ig heavy chain, se
118      257.5     9.5      342      2      A45966      Ig alpha chain C r
119      257      9.5      111      2      S43148      Ig upatlon chain -
120      253.5     9.4      573      2      S12838      Ig mu chain precur
121      245      9.1      335      2      S09275      Ig alpha chain C r
122      243.5     9.0      220      2      C22360      Ig alpha-2 chain C
123      241.5     8.9      448      2      A46533      Ig heavy chain C r
124      223.5     8.3      402      2      S20002      Ig heavy chain, se
125      223.5     8.3      1005     2      T18537      Ig heavy chain - C

```

ALIGNMENTS

RESULT 1

GHMU Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #Sequence revision 18-Aug-1982 #text change 16-Jul-1999

C:Accession: A53433; #Accession: S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A53433; MUID:82274238; PMID:6287432

A:Accession: A53433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>

A:Cross-references: EMBL:Z17370

A>Note: this sequence has the Gln(17) allotypic marker, 97-Lys, and the Gln(1) markers,

R:Harrie, L.J.

submitted to the EMBL Data Library, October 1992

A:Reference number: S33904

A:Accession: S36861

A:Molecule type: DNA

A:Residues: 2-330 <HAR>

A:Cross-references: EMBL:Z17370

R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaide, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of a

A:Reference number: S33887; MUID:83001943; PMID:6811139

A:Accession: S33887

A:Molecule type: DNA

A:Residues: 88-113:235-330 <TAK>

A:Cross-references: EMBL:Z17370

R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,

Biochemistry 9, 3161-3170, 1970

A:Title: The covalent structure of a human gammag-1 immunoglobulin. VII. Amino acid sequen

A:Reference number: A90563; MUID:71064024; PMID:5489771

A:Contents: myeloma protein Eu

A:Accession: B90563

A:Molecule type: protein

A:Residues: 1-96, 'R', '98-135 <CON>

A>Note: this sequence has the Gln(3) marker, 97-Arg

R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A:Title: The covalent structure of a human gammag-1 immunoglobulin. VIII. Amino acid sequen

A:Reference number: A90564; MUID:71064025; PMID:5530842

A:Contents: Eu

```

A:Accession: A90564
A:Molecule type: protein
A:Residues: 136-154, 'Q', 156-165, 'Q', 167-176, 'Q', 178-194, 'N', 196-197, 'D', 199-238, 'E', 240,
A>Note: this sequence has the Gln(non-1) markers, 239-Glu and 241-Met
R:Pomstingl, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A:Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie),
igen Primärstruktur.
A:Reference number: A91668; MUID:77070269; PMID:826475
A:Contents: myeloma protein Nie
A:Accession: B91668
A:Molecule type: protein
A:Residues: 1-34, 'Q', 36-96, 'K', 98-115, 'Q', 117-197, 'D', 199-238, 'D', 240, 'L', 242-268, 'E', 27,
A>Note: this sequence has the Gln(17) and Gln(1) markers
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A:Title: Die Primärstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOL
A:Reference number: A91723; MUID:8328131; PMID:6884994
A:Accession: A91723
A:Molecule type: protein
A:Residues: 1-96, 'R', 98-197, 'D', 199-238, 'E', 240, 'W', 242-266, 'D', 268-271, 'D', 273-330 <SCH
A>Note: this sequence has the Gln(3) and Gln(non-1) markers
R:Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A:Title: The covalent structure of a human gammag-1 immunoglobulin. X. Intrachain disulfide
A:Reference number: A90565; MUID:71064027; PMID:4923144
A:Contents: annotation; disulfide bonds
R:Drexler, L.; Schwarz, J.; Reiche, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
endromide cleavage products, and the disulfide bridges.
A:Reference number: A91667; MUID:77070267; PMID:1002129
A:Contents: annotation; disulfide bonds
C:Genetics:
A:Gene: GDB:IGHG1
A:Cross-references: GDB:120085; OMIM:147100
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 114/1; 224/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa)
chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la
C:Superfamily: Immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:27-83,144-204,250-308/Disulfide bonds: #status experimental
F:103/Disulfide bonds: interchain (to light chain) #status experimental
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

```

```

Query Match 45.4%; Score 1227.5; DB 1; Length 330;
Basic Local Similarity 74.5%; Pred. No. 7.1e-66;
Matches 246; Conservative 7; Mismatches 22; Indels 55; Gaps 6;

QY 151 PSVQCRSPRGKNIQGG-----TLTSVS-----QLELDPSG-----180
DQ 6 PSVFPPLAPSSKSTSGGALGCLVKQVFPPEPVYVSNMKGALTSQVHFFPAVLQSSGLYSL 65
DQ 181 -----TWTCVLNQNQKVEFKIDIVPCPAPPKSCDKTHTC-----PELLG 221
DQ 66 SSVVTVPVSSSLGTQTYICNV---NHKPSNTKVD-----KKVFPKSCDKTHTCPCPAPPELLG 119
QY 222 GPSVFLPPPKPKXTMISRTPEVTCVVVDVSHEDPEVKFMWYVDGVVHNAKTKRPREQY 261
DQ 120 GPSVFLPPPKPKXTMISRTPEVTCVVVDVSHEDPEVKFMWYVDGVVHNAKTKRPREQY 179
QY 282 NSTYRVSVTVTVAHQDLINKEKYCKVSNKALPAPIEKTSKAGQPREPOVYTLPPSRD 341
DQ 180 NSTYRVSVTVTVAHQDLINKEKYCKVSNKALPAPIEKTSKAGQPREPOVYTLPPSRD 239
QY 342 ELTKQVSLTCLVKGFPYSDIAVESNGGPENNYKTPPVLDGSGFPFLYSKLTVDKSR 401

```


Db 240 ELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTPTPLVDSGDFLYSKLTIVDKSR 299
QY 402 WQGNVFGSCVMHEALHNNHYTOKSLSLSPG 431
Db 300 WQGNVFGSCVMHEALHNNHYTOKSLSLSPG 329

RESULT 2

569339
Ig heavy chain V region precursor - human

CISpecies: Homo sapiens (man)
CISDate: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
CISAccession: S69339; S72664
R:Khamlich, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995

A>Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A:Reference number: S69339; PMID:95262687; PMID:7744049
A:Accession: S69339

A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
R:Khamlich, A.A.

A:Submitted to the EMBL Data Library, September 1994
A:Reference number: S72664
A:Accession: S72664

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-140, 'C', 142-374 <KH2>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 45.2%; Score 1221; DB 2; Length 374;
Best Local Similarity 71.4%; Pred. No. 2e-65;
Matches 242; Conservative 19; Mismatches 34; Indels 44; Gaps 5;

QY 137 QSLTTLT-----LESPGSS-----SPSVQCRSPRGKNI 163
Db 35 QTLTLCTSSGSLSSGCVGCKIRPPGQALEMLALITWDDDKRSPSLRRLTITTKOT 94
QY 164 QGKGLTSLVSOELQDSGTMTCTVL-----QNKVVEFKIDIVPCPAPRPKSCDKTHTC- 216
Db 95 SKNQVLTMTNVDPAATATAYGSGYVEGGYGRFHSWQGLTYVYSSSPKSGDKHTHP 154
QY 217 ----PELLGSPSVFLFPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFMYVDGVEVNA 272
Db 155 PCPAPPELLGSPSVFLFPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFMYVDGVEVNA 214
QY 273 KTKPREQVNSTYRVVSVLTVLHODWLNGEKYCKVSNKALPAPLEKTSKAKGQPREPQ 332
Db 215 KTKPREQVNSTYRVVSVLTVLHODWLNGEKYCKVSNKALPAPLEKTSKAKGQPREPQ 274
QY 333 VTTLPSPRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTPTPLVDSGDFLY 392
Db 275 VTTLPSPRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTPTPLVDSGDFLY 334
QY 393 SKLTVDKSRWQGNVFGSCVMHEALHNNHYTOKSLSLSPG 431
Db 335 SKLTVDKSRWQGNVFGSCVMHEALHNNHYTOKSLSLSPG 373

RESULT 3

531866

Ig gamma-1 chain C region - synthetic

CISpecies: synthetic

A>Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli

CISDate: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000

CISAccession: S31866

R:Filipula, D.

submitted to the EMBL Data Library, February 1993

A>Description: Screening method for protein-protein interactions of cloned gene products.
A:Reference number: S31866
A:Accession: S31866

A:Molecule type: mRNA
A:Residues: 1-255 <FIL>
A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069
C:Keywords: immunoglobulin
F:1-22/Region: Escherichia coli outer membrane protein A precursor
F:23-255/Region: human Ig gamma-1 chain C region

Query Match 44.5%; Score 1203.5; DB 4; Length 255;
Best Local Similarity 97.4%; Pred. No. 1.4e-64;
Matches 225; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 206 EPKSCDKTHTC-----PELLGSPSVFLFPKPKDTLMSRTPEVTCVVVDVSHEDPEVK 260
Db 24 ESKSCDKHTHTCPCPAPPELLGSPSVFLFPKPKDTLMSRTPEVTCVVVDVSHEDPEVK 83
QY 261 NMYVDGVEVNAKTKPREQVNSTYRVVSVLTVLHODWLNGEKYCKVSNKALPAPLEK 320
Db 84 NMYVDGVEVNAKTKPREQVNSTYRVVSVLTVLHODWLNGEKYCKVSNKALPAPLEK 143
QY 321 ISKAGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTT 380
Db 144 ISKAGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTT 203
QY 381 PVLDSGDFLYSKLTVDKSRWQGNVFGSCVMHEALHNNHYTOKSLSLSPG 431
Db 204 PVLDSGDFLYSKLTVDKSRWQGNVFGSCVMHEALHNNHYTOKSLSLSPG 254

RESULT 4

PT0207

Ig gamma chain C region - chimpanzee

CISpecies: Pan troglodytes (chimpanzee)

CISDate: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999

CISAccession: PT0207

R: Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.

Mol. Immunol. 28, 319-322, 1991

A>Title: Nucleotide sequence of chimpanzee Fc and hinge regions.

A:Reference number: PT0207; PMID:91287716; PMID:2062315

A:Accession: PT0207

A:Molecule type: mRNA

A:Residues: 1-234 <EHR>

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 43.4%; Score 1173.5; DB 2; Length 234;
Best Local Similarity 92.5%; Pred. No. 7.6e-63;
Matches 221; Conservative 3; Mismatches 2; Indels 13; Gaps 2;

QY 192 KVEFKIDIVPCPAPRPKSCDKTHTC-----PELLGSPSVFLFPKPKDTLMSRTPEVTC 246
Db 4 KVDKIV-----EPKSCDTHTCPCAPPELLGSPSVFLFPKPKDTLMSRTPEVTC 55
QY 247 VVVDVSHEDPEVKFMYVDGVEVNAKTKPREQVNSTYRVVSVLTVLHODWLNGEKYCK 306
Db 56 VVVDVSHEDPEVKFMYVDGVEVNAKTKPREQVNSTYRVVSVLTVLHODWLNGEKYCK 115
QY 307 KVSNNKALPAPLEKTSKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEM 366
Db 116 KVSNNKALPAPLEKTSKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEM 175
QY 367 ESNQGPENNYKTTPTPLVDSGDFLYSKLTVDKSRWQGNVFGSCVMHEALHNNHYTOKS 425
Db 176 ESNQGPENNYKTTPTPLVDSGDFLYSKLTVDKSRWQGNVFGSCVMHEALHNNHYTOKS 234

RESULT 5

A23511

Ig gamma-3 chain C region (allotype G3m(b)) - human

CISpecies: Homo sapiens (man)

CISDate: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999

CISAccession: A23511

R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.

A>Note: this sequence may represent an allelic form or another gamma chain subclass
C:Comment: The heavy chain disease protein Wis is shown.

C:Genetics:

A:Gene: GDB:IGHG3

A:Cross-references: GDB:119339; OMIM:147120

A:Map position: 14q32.33-14q32.33

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglyutamic acid

F:203-270/Domain: immunoglobulin homology <IMM>

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 41.9%; Score 1133; DB 1; Length 289;

Best Local Similarity 88.2%; Pred. No. 2,5e-60;

Matches 209; Conservative 13; Mismatches 9; Indels 6; Gaps 2;

QY 201 PCP-ADPKSCDKTHTC-----PELIGPSVFLPPKPDITMISTPEVTCVVDVSH 254

DB 53 PCPRCPKPCSCDTPPCPCPCPAPPELLIGPSVFLFPKPDITMISTPEVTCVVDVSH 112

QY 255 DEVRKKNVVDGVEVNAATKPREQVNSTYRVVSVTLTHODMLNGKRYCKVSNKALP 314

DB 113 DEVRKKNVVDGVEVNAATKPREQVNSTYRVVSVTLTHODMLNGKRYCKVSNKALP 172

QY 315 APIEKTISKAKGPREPQVYTLPPSRDELTKNQVSLTCLVKGFPSDIAVEMESNQPN 374

DB 173 APIEKTISKAKGPREPQVYTLPPSRDELTKNQVSLTCLVKGFPSDIAVEMESNQPN 232

QY 375 NYKTPPVLDSDGFFLYSKLTVDKSRMQQGVNFSQSVNHEALHNHYTQKSLSLSPG 431

DB 233 NYKTPPVLDSDGFFLYSKLTVDKSRMQQGVNFSQSVNHEALHNHYTQKSLSLSPG 289

RESULT 8

IG gamma-2 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 30-Apr-1981 #sequence revision 13-Jun-1983 #text_change 21-Jul-2000

C:Accession: A93906; A92809; A90752; A93132; A02148

R:Billson, J.; Hood, L.

Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982

A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con

A:Reference number: A93906; MUID:82197621; PMID:6804948

A:Accession: A93906

A:Molecule type: DNA

A:Residues: 1-326 <EL>

A:Cross-references: GB:J00230; NID:932759; PIDN:CABS8438.1; PID:9606056

A>Note: Lys-326 is probably removed posttranslationally

R:Wang, A.C.; Tung, E.; Fudenberg, H.H.

J. Immunol. 125, 1048-1054, 1980

A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and E

A:Reference number: A92809; MUID:81007873; PMID:6774012

A:Accession: A92809

A:Contents: myeloma protein t11

A:Molecule type: protein

A:Residues: 1-119, 'Q', '21-57', 'Z', '59', 'A', '61-193', 'D', '195-325 <MAN>

A>Note: Trp-156 is at or near the complement-binding site

R:Connell, G.E.; Parr, D.M.; Hofmann, T.

Can. J. Biochem. 57, 758-767, 1979

A:Title: The amino acid sequences of the three heavy chain constant region domains of a

A:Reference number: A90752; MUID:80001357; PMID:113060

A:Accession: A90752

A:Contents: myeloma protein z1e

A:Molecule type: protein

A:Residues: 1-24, 'E', '26-57', 'EV', '60-85', '132-171', 'ZZZ', '175', 'B', '177-193', 'D', '195-196', 'Q', '198-

A>Note: this sequence has since been revised

R:Hofmann, T.; Parr, D.M.

Mol. Immunol. 16, 923-925, 1979

A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g

A:Reference number: A93132; MUID:80114419; PMID:118920

A:Accession: A93132

A:Contents: z1e

A:Molecule type: protein

A:Residues: 238-275 <HOF>
R:Hofmann, T.; Parr, D.M.
submitted to the Atlas, March 1980
A:Reference number: A94591
A:Contents: annotation; z1e, revisions to residues 25, 59, 60, and 264-268
A>Note: the revised sequence differs from that shown in having 60-Ala and in the amidat

ned

R:Milstein, C.; Frangione, B.

Biochem. J. 121, 217-225, 1971

A:Title: Disulfide bridges of the heavy chain of human immunoglobulin G2.

A:Reference number: A90253; MUID:72033500; PMID:4940472

A:Contents: annotation; myeloma protein Sa, disulfide bonds

R:Frangione, B.; Milstein, C.; Pink, J.R.L.

Nature 221, 145-148, 1969

A:Title: Structural studies of immunoglobulin G.

A:Reference number: A93157; MUID:69064124; PMID:5782707

A:Contents: annotation; Sa, disulfide bonds

C:Genetics:

A:Gene: GDB:IGHG2

A:Cross-references: GDB:119338; OMIM:147110

A:Map position: 14q32.33-14q32.33

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka

tain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:133-202/Domain: immunoglobulin homology <IM1>

F:133-306/Domain: immunoglobulin homology <IM2>

F:14/Disulfide bonds: interchain (to light chain) #status experimental

F:27-93,140-200,246-304/Disulfide bonds: #status experimental

F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental

F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.8%; Score 1130.5; DB 1; Length 326;

Best Local Similarity 82.6%; Pred. No. 4e-60;

Matches 214; Conservative 12; Mismatches 12; Indels 21; Gaps 3;

QY 181 TWTCV-----LQNKVFEK-IDIVCPAPPEKSCDKTCEBLGSPVFLPPKP 232

DB 80 TYTCVNDHKPSNTKDKVTKKVCCEPCPCP-----VAGPSVFLPPKP 126

QY 233 KDTLMISRTPEVTCVVDVSHEDPEVKFNMYVDGVEVNAATKPREQVNSTYRVVSVLT 292

DB 127 KDTLMISRTPEVTCVVDVSHEDPEVKFNMYVDGVEVNAATKPREQVNSTYRVVSVLT 186

QY 293 VHQDMLNGKEYCKVSNKALPAPIEKTISKAKGPREPQVYTLPPSRDELTKNOVSLTC 352

DB 187 VHQDMLNGKEYCKVSNKALPAPIEKTISKAKGPREPQVYTLPPSRDELTKNOVSLTC 246

QY 353 LVKGFYPSDIAVEMESNQPNKYTPPVLDSDGFFLYSKLTVDKSRMQQGVNFSQSV 412

DB 247 LVKGFYPSDIAVEMESNQPNKYTPPVLDSDGFFLYSKLTVDKSRMQQGVNFSQSV 306

QY 413 MHEALHNHYTQKSLSLSPG 431

DB 307 MHEALHNHYTQKSLSLSPG 325

RESULT 9

GAHU

IG gamma-4 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 02-Apr-1982 #sequence revision 02-Apr-1982 #text_change 16-Jul-1999

C:Accession: A90933; A90249; A02150

R:Billson, J.; Hood, L.

DNA 1, 11-18, 1981

A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.

A:Reference number: A90933; MUID:83157104; PMID:6299662

A:Accession: A90933

A:Molecule type: DNA

A:Residues: 1-327 <EL>

A>Note: the sequence was determined from the germline gene

R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.

Biochem. J. 117, 33-47, 1970

A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constant

A:Reference number: A90249; MUID:70207560; PMID:4192699

A:Accession: A90249

A:Molecule type: protein

A:Residues: 1-30/81-326 <PIN>

C:Genetics:

A:Gene: GDB:IGHG4

A:Cross-references: GDB:119340; OMIM:147130

A:Map position: 14q32.33-14q32.33

A:Interons: 99/1, 111/1, 221/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa)

chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la

C:Superfamily: immunoglobulin C region, immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:20-85/Domain: immunoglobulin homology <IM1>

F:99-110/Region: hinge

F:134-203/Domain: immunoglobulin homology <IM3>

F:240-307/Domain: immunoglobulin homology <IM2>

F:14/Disulfide bonds: interchain (to light chain) #status experimental

F:127-83,141-201,247-305/Disulfide bonds: #status predicted

F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental

F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.4%; Score 1119; DB 1; Length 327;

Best Local Similarity 83.7%; Pred. No. 1.9e-59;

Matches 215; Conservative 11; Mismatches 15; Indels 16; Gaps 4;

QY 181 TWCTV-----LQKKVKEFKIDVPCPAEPKSCDHTHCPELGGPSVFLPPPKD 234

DB 80 TYTCNVDPKPSNTKYDKRVESKYG-PPCP-----SC-----PAPEFLGGPSVFLPPPKD 129

QY 235 TLMISRPETVCVVVDVSHEDPEVKFNMYVDGVEVNAKTKPREQNSTYRVSVLTVL 294

DB 130 TLMISRPETVCVVVDVSHEDPEVKFNMYVDGVEVNAKTKPREQNSTYRVSVLTVL 189

QY 295 HODWLNGEYKCKSNKALPAPIEKTISKAGOPREPQVYTLPPSRDELTKNOVSLTCLV 354

DB 190 HODWLNGEYKCKSNKALPAPIEKTISKAGOPREPQVYTLPPSRDELTKNOVSLTCLV 249

QY 355 KGFPYPSDIAVWESNGQENNYKTPPYLSDGDSFFLYSKLTYPKSRQGNVSCSYMH 414

DB 250 KGFPYPSDIAVWESNGQENNYKTPPYLSDGDSFFLYSKLTYPKSRQGNVSCSYMH 309

QY 415 EALHNHYTKSLSLSPG 431

DB 310 EALHNHYTKSLSLSLG 326

RESULT 10

RHWU74

T-cell surface glycoprotein CD4 precursor [validated] - human

N:Alternate names: T-cell surface antigen T4/Lew 3

C:Species: Homo sapiens (man)

C:Date: 28-May-1986 #sequence_revision 31-Dec-1986 #text_change 20-Apr-2001

C:Accession: A90872; A32722; A34194; A53287; I54176; I54297; A02109; A30039

R:Maddon, P.J.; Littman, D.R.; Godfrey, M.; Maddon, D.E.; Chess, L.; Axel, R.

Cell 42, 93-104, 1985

A:Title: The isolation and nucleotide sequence of a cDNA encoding the T cell surface pro

A:Accession: A90872; MUID:85254948; PMID:2990730

A:Accession: A90872

A:Molecule type: mRNA

A:Residues: 1-25, 'N', 27-458 <MAD>

A:Experimental source: clone pT4B

R:Littman, D.R.; Maddon, P.J.; Axel, R.

Cell 55, 541, 1988

A:Title: Corrected CD4 sequence.

A:Reference number: A90907; MUID:89028665; PMID:3262123

A:Contents: annotation; revision to residue 26

R:Cammerini, D.; Seed, B.

Cell 60, 747-754, 1990

Article: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi

A:Reference number: A32722; MUID:90182664; PMID:2107024

A:Accession: A32722
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 26-426, 428-458 <CAM>
R:Car, S.A.; Hemling, M.E.; Folana-Wasseman, G.; Sweet, R.W.; Anumula, K.; Barr, J.R.; J. Biol. Chem. 264, 21286-21295, 1989
A>Title: Protein and carbohydrate structural analysis of a recombinant soluble CD4 receptor
A:Reference number: A3194; PMID:90078232; PMID:2552374
A:Contents: disulfide bonds; carbohydrate-binding sites
A:Accession: A34194
A:Molecule type: protein
A:Residues: 26-394 <CAR>
R:Leeman, S.; DeMartino, J.A.; Daugherty, B.L.; Foeldvari, I.; Yellin, M.J.; Cleary, A.
A>Title: A single amino acid substitution in a common African allele of the CD4 molecule
A:Reference number: A53287; PMID:9202595; PMID:1961196
A:Accession: A53287
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 250-264, 'W', 266-280 <LED>
A>Note: sequence extracted from NCBI backbone (NCBIP:68249)
R:Edwards, M.C.; Gibbs, R.A. Genomics 14, 590-597, 1992
A>Title: A human dimorphism resulting from loss of an Alu.
A:Reference number: I54176; PMID:93052387; PMID:1330888
A:Accession: I54176
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-72 <RSS>
A:Cross-references: GB:U47924; GB:M6525; GB:U72506; NID:91633547; PIDN:AA51309.1; PID: R.Hodge, T.W.; Sasse, D.R.; McDougal, J.S. Hum. Immunol. 30, 99-104, 1991
A>Title: Humans with OKT4-epitope deficiency have a single nucleotide base change in the A:Reference number: I54297; PMID:91216786; PMID:1708753
A:Accession: I54297
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-264, 'W', 266-458 <RE2>
A:Cross-references: GB:M3160; NID:9179143; PIDN:AAA6069.1; PID:9179144
C:Comment: Macrophage tropic strains of HIV-1 bind to a complex of chemokine (C-C) receptor C:Genetics:
A:Gene: GDB:CD4
A:Cross-references: GDB:119767; OMIM:186940
A:Map position: 12pter-12p12
A:Intron: 16/3
C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
C:Keywords: AIDS; duplication; glycoprotein; T-cell; transmembrane protein
F:1-35/Domain: signal sequence #status predicted <SIG>
F:26-458/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>
F:34-111/Domain: immunoglobulin homology <IM1>
F:136-166/Domain: immunoglobulin homology #status atypical <IM2>
F:216-299/Domain: immunoglobulin homology <IM3>
F:321-372/Domain: immunoglobulin homology <IM4>
F:397-426/Domain: transmembrane #status predicted <TM>
F:421-458/Domain: intracellular #status predicted <IN>
F:41-109, 155-184, 328-370/Disulfide bonds: #status experimental
F:296,325/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match	37.9%	Score 1023	DB 1	Length 458
Best Local Similarity	99.5%	Pred. No. 1.5e-53		
Matches 199	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY 1	MNRGVPFRLILLVLQALIPATQGNKVVGGKGDVETLCTASQKSIQPHMKNNDIK 60			
DB 1	MNRGVPFRLILLVLQALIPATQGNKVVGGKGDVETLCTASQKSIQPHMKNNDIK 60			
QY 61	ILGNQGSFLTKGPSKINDRADSRSIMDQGNFPIINLKIEDSDTYICEVEDQKEEVQL 120			
DB 61	ILGNQGSFLTKGPSKINDRADSRSIMDQGNFPIINLKIEDSDTYICEVEDQKEEVQL 120			
QY 121	LVFGITANSDFHLLQGSITLTIESPPGSSPVQCRPRGNIQGGKTLVSQLELDQSG 180			
DB 121	LVFGITANSDFHLLQGSITLTIESPPGSSPVQCRPRGNIQGGKTLVSQLELDQSG 180			

QY 181 TWCTVLONQKKEFKIDIV 200
 |||||
 DB 181 TWCTVLONQKKEFKIDIV 200

RESULT 11

G3MSM
 Ig gamma-3 chain C region, membrane-bound form - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 13-Aug-1986 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C/Accession: A02156; A02155
 R/Author: J.A.; Word, C.J.; Rimm, D.; Der-Balan, G.P.; Martinez, H.M.; Tucker, P.W.; Blach
 EMO J. 3, 2041-2046, 1984
 A/Title: Structure analysis of the murine IgG3 constant region gene.
 A/Reference number: A02156; MUID:85027161; PMID:6092053
 A/Accession: A02156

A/Molecule type: DNA

A/Residues: 1-398 <MBL>

A/Cross-references: GB:J00451; NID:g194392; PIDN:AAB59655.1; PID:g194433

A/Note: The sequence was determined from the germline gene

R/Komarov, M.; Clayton, L.; Rogers, J.; Robertson, S.; Keltman, J.; Wall, R.

Nucleic Acids Res. 11, 6775-6785, 1983

A/Title: The structure of the mouse immunoglobulin in gamma-3 membrane gene segment.

A/Reference number: A02155; MUID:84041483; PMID:6314258

A/Accession: A02155

A/Molecule type: DNA

A/Residues: 328-332, 'G', 334-341, 'Q', 343-387, 'F', 389-398 <KOM>

A/Cross-references: GB:K00688

A/Note: the sequence was determined from the germline gene

C/Genetics:

A/Intons: 97/1; 113/1; 223/1; 328/1; 371/3

C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (L) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 12

C/Superfamily: Immunoglobulin C region; immunoglobulin homology

C/Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob

F/19-83/Domain: immunoglobulin homology <IM1>

F/97-112/Region: hinge

F/136-205/Domain: immunoglobulin homology <IM2>

F/242-309/Domain: immunoglobulin homology <IM3>

F/346-362/Domain: transmembrane #status predicted <TM>

F/363-398/Domain: intracellular #status predicted <INT>

F/179,322/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 35.4%; Score 955.5; DB 1; Length 398;

Best Local Similarity 61.4%; Pred. No. 1,3e-49;

Matches 180; Conservative 40; Mismatches 58; Indels 15; Gaps 4;

QY 170 SVSLELDSDGTV-----TCTVLONQKKEFKIDI---VPCAPPEPKSCDKHTCP--EL 219

DB 62 SLSSLVTVSSTYSPSCQVTCNVANPASKTELRIEPRIPKSTPPGS-----SCPPGNI 116

QY 220 LGSPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDSGEVFNAAKTPRRE 279

DB 117 LGSPSVFIFPPKPKDMLMISLTPEKTCVVVDVSEDDPDVAVSVFVNKKEVHTAMTPRRA 176

QY 280 QVNSTRVSVTLVTLQDMLNGKEYCKRYSKALLPAPIETIKISKAGQPREPVYTLPPS 339

DB 177 QVNSTRVSVTLVTLQDMLNGKEYCKRYSKALLPAPIETIKISKAGQPREPVYTLPPS 236

QY 340 RDELTKNOVSLTLGVGFSYSDIAVEMESNGDEPNYKTPPYLSDGSEFVLSKLTIVK 399

DB 237 RDELTKNOVSLTLGVGFSYSDIAVEMESNGDEPNYKTPPYLSDGSEFVLSKLTIVK 296

QY 400 SRNQGNVFSQVMEALNHYTQKSLSPGLDDETCADGDELGLMTT 452

DB 297 DSVLQGEIFTCQVVEALNHYTQKSLSPGLDDETCADGDELGLMTT 349

RESULT 12

G1MSM
 Ig gamma-1 chain C region, membrane-bound form - mouse
 C/Species: Mus musculus (house mouse)

C/Date: 17-Dec-1982 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999

C/Accession: B02159; A02160; B02158

R/Honjo, T.; Obata, M.; Yamawaki-Kataoka, Y.; Kataoka, T.; Kawakami, T.; Takahashi, N.;

Cell 18, 559-568, 1979

A/Title: Cloning and complete nucleotide sequence of mouse immunoglobulin gamma1 chain

A/Reference number: A02159; MUID:80045036; PMID:115593

A/Accession: B02159

A/Molecule type: DNA

A/Residues: 1-393 <HON>

A/Cross-references: GB:J00453

A/Note: the sequence was determined from the germline gene

R/Tyler, B.M.; Cowan, A.F.; Gerondakis, S.D.; Adams, J.M.; Bernard, O.

Proc. Natl. Acad. Sci. U.S.A. 79, 2008-2012, 1982

A/Title: mRNA for surface immunoglobulin gamma chains encodes a highly conserved transm

A/Reference number: A02160; MUID:82197626; PMID:6804950

A/Accession: A02160

A/Molecule type: mRNA

A/Residues: 323-393 <TYL>

R/Rogers, J.; Choi, E.; Souza, L.; Carter, C.; Word, C.; Kuehl, M.; Bisenberg, D.; Wall,

Cell 26, 19-27, 1981

A/Title: Gene segments encoding transmembrane carboxyl termini of immunoglobulin gamma

A/Reference number: A02158; MUID:82115295; PMID:6799207

A/Accession: B02158

A/Molecule type: DNA

A/Residues: 323-366 <ROG>

A/Note: this sequence is the translation of the first exon of the M segment

C/Comment: Cell lines producing IgG contain two mRNA species for Ig gamma chains. The m

code membrane-bound chains in that it contains an alternative 3' end, encoded in separa

C/Genetics:

A/Intons: 1/1; 98/1; 111/1; 218/1; 323/1; 366/3

C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (L) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 12

C/Superfamily: Immunoglobulin C region; immunoglobulin homology

C/Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob

F/131-200/Domain: immunoglobulin homology <IM1>

F/240-357/Domain: transmembrane #status predicted <TM>

F/358-393/Domain: intracellular #status predicted <INT>

F/174,278/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 34.5%; Score 932.5; DB 1; Length 393;

Best Local Similarity 55.0%; Pred. No. 2,9e-48;

Matches 177; Conservative 50; Mismatches 60; Indels 35; Gaps 5;

QY 133 LQGSQSLTLTSPSSSPVOCSPRGKNIQGGKTLVSQLELDSDGTVTCVLO--NQ 190

DB 56 VQGSQSLTLTSSSVTVSSP-----RSEVTCNVANPASP 90

QY 191 KYVEFKIDIVPCPAPPEPKSCDKHTCPPELLGSPVFLPPPKDPTLMISRTPEVTCVVD 250

DB 91 TVYDKKIVRDGCG-KPCIC-----TVPEV---SSVFIFPPKPKDVLTLTLTKVTCVVD 142

QY 251 VSHEDPEVKFNNYVDSGEVFNAAKTPREQVNSTRVSVTLVTLQDMLNGKEYCKRYAN 310

DB 143 ISKDPEVQFSWFDVDEVTATQTOREQFNSFVSVELPIHMDMLNGKEYCKRYAN 202

QY 311 KALPAIETIKISKAGQPREPVYTLPPSRDRLTKNOVSLTLGVGFSYSDIAVEMESNG 370

DB 203 AAPPAIETIKISKAGQPREPVYTLPPSRDRLTKNOVSLTLGVGFSYSDIAVEMESNG 262

QY 371 QPENNYKTPPYLSDGSEFVLSKLTIVDSRNQGNVFSQVMEALNHYTQKSLSP 430

DB 263 QPAENYKTPPYLSDGSEFVLSKLTIVDSRNQGNVFSQVMEALNHYTQKSLSP 322

QY 431 GLQDDETCADGDELGLMTT 452

DB 323 GLQDDETCADGDELGLMTT 344

RESULT 13

S22080
 Ig heavy chain precursor (B/MT.4A.17.H5.A5) - bovine
 N/Alternate names: Ig gamma-1 chain C region (clone 8.10)
 C/Species: Bos primigenius taurus (cattle)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C/Accession: S32080; S06610; A31303
R/Sanders, P.G.
Submitted to the EMBL Data Library, November 1991
A/Reference number: S32080
A/Accession: S32080
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-470 <SAS>
A/Cross-references: EMBL:X62916; NID:q439; PIDN:CAA44699.1; PID:q440
R/Symons, D.B.A.; Clarkson, C.A.; Beale, D.
Mol. Immunol. 26, 841-850, 1989
A/Title: Structure of bovine immunoglobulin constant region heavy chain gamma 1 and gamma 2
A/Reference number: S06610; MUID:90097956; PMID:2513487
A/Accession: S06610
A/Molecule type: DNA
A/Residues: 142-470 <SYM>
A/Cross-references: EMBL:X16701
A/Note: The sequence was determined from the germline gene
C/Genetics:
A/Genes: Ig CH gamma-1
A/Introns: 98/1; 111/1; 221/1
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: glycoprotein; heterotrimer; immunoglobulin; membrane protein
F/161-225/Domain: immunoglobulin homology <IMM>
F/118/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 33.9%; Score 915.5; DB 2; Length 470;
Best Local Similarity 43.6%; Pred. No. 3.7e-47;
Matches 214; Conservative 52; Mismatches 136; Indels 89; Gaps 15;

QY 6 PFRHLTLVQ-----LALLPATQGNKVLGKGDVLTCTAS--QKSKIQFMKNSNQ 58
DB 3 PLMTLLFLVLSAPIGVLSQVQRESGPLV--KPSQTLSLTCTVSGFSLSSALTMVRAP 60
QY 59 IKILGNQGSFLTKGSKLANDRDSRRSLMDQCN--FLIIRKMLKIEISDYIC----- 109
DB 61 GKALEWVGITSGGTTYYNPAKSLRSLTKENSKSQVLSVSVTPBEDTATYCCARSTYG 120
QY 110 EVEDQKEVQVLVFLTANSPTHLILQGSLLTLTLESPPGSSPSVQ-----CRSPRGKNIQ 164
DB 121 EVGD-----GALIDA-----WGQGLVTVSSASTAPKVPYPLSSCCGDKSSSTVT 165
QY 165 GSKTLVSQLE-----LQDSG-----TWLCTV 186
DB 166 LGCLVSVMPPEPVYTNMNGALKSGVHTFPAVLQSSGLYSLSMTVTVGTSQGITCNV 225
QY 187 LQ--NOKKVERKIDIVPCPAPBPXSCDKTHTC--PELLGSPVLEFPKXPOTLMISKTP 242
DB 226 AHPASSTKVDKAVD---PTCKPSPCD---CCPPELPFGSPSVTFPPKPMOTLLISGTP 278
QY 243 EVTCVVVDVSHEDPEVKFNWYVDGVEVNAKTKPREEQYNSTYRVSVLTALHDDMLNGK 302
DB 279 EVTCVVVDVGHDDPEVKFSWFDVDEVTATTPKEEDFQNSTYRVASLRQIHDDMTGSK 338
QY 303 EYCKVSNKALPAPIEKTIKSAKQPRPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 362
DB 339 EFKCVNHEGLPAPIVTRTISRTKGPAREPQVYVLAPOEELSKGVSITCMVTSFYPIYI 398
QY 363 AVENESNQP--ENNYKTTTPVLDSGSFYLKSLTVDKSRWQGNVFSVCMEDALHNH 420
DB 399 AVENQNRNQPSSEDEKYGTTTPQLDADSSYFLYSKLRVDNRNMQEGDVTYTCVVMEDALHNH 458
QY 421 YTKSLSLSPG 431
DB 459 YTKSTSKSAG 469

RESULT 14
GZMSAM
Ig gamma-2a chain C region, membrane-bound form - mouse
C/Species: Mus musculus (house mouse)
C/Date: 19-Feb-1984 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999

C/Accession: A02154; B32657; I57809
R/Yanawaki, Katakoka, Y.; Nakai, S.; Miyata, T.; Honjo, T.
Proc. Natl. Acad. Sci. U.S.A. 79, 2623-2627, 1982
A/Title: Nucleotide sequences of gene segments encoding membrane domains of immunoglobulin A
A/Reference number: A02154; MUID:8222190; PMID:6283537
A/Accession: A02154
A/Molecule type: DNA
A/Residues: 329-399 <YAM>
A/Cross-references: GB:J00471
A/Note: The sequence was determined from the germline gene
R/Yanawaki, Katakoka, Y.; Miyata, T.; Honjo, T.
Nucleic Acids Res. 9, 1365-1381, 1981
A/Title: The complete nucleotide sequence of mouse immunoglobulin gamma 2 A gene and epsilon
A/Reference number: A32657; MUID:81198976; PMID:6262729
A/Accession: B32657
A/Molecule type: DNA
A/Residues: 1-329 'K' <YAN>
R/Hall, B.; Micalarek, C.
Mol. Immunol. 26, 819-826, 1989
A/Title: Sequence and polyadenylation site determination of the murine immunoglobulin gamma 2A gene
A/Reference number: I57809; MUID:90097953; PMID:2513466
A/Accession: I57809
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 373-399 <RES>
A/Cross-references: GB:M35032; NID:q194478; PIDN:AAA37919.1; PID:q387217
C/Comment: The sequence of residues 1-328 was assumed to be identical with the corresponding
C/Comment: Cell lines producing IgG contain two mRNA species for Ig gamma chains. The major
C/Comment: hat it contains an alternative 3' end, encoded in separate exons, that is homologous with
C/Genetics:
A/Introns: 1/1; 98/1; 114/1; 224/1; 329/1; 372/1
C/Complex: An immunoglobulin heterotrimer subunit consists of two identical light (Kap
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: alternative splicing; duplication; glycoprotein; heterotrimer; immunoglobulin
F/137-206/Domain: immunoglobulin homology <IMM>
F/346-363/Domain: transmembrane #status predicted <TM>
F/364-399/Domain: intracellular #status predicted <INT>
F/180/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 33.5%; Score 906.5; DB 1; Length 399;
Best Local Similarity 60.5%; Pred. No. 1e-46;
Matches 173; Conservative 34; Mismatches 60; Indels 19; Gaps 5;

QY 179 SGTW-----TCTVQ--NOKKVERKID-----IYPCPAPBPXSCDKTHTCPELLGSPSVF 226
DB 72 SSTWPSQSTICNVNAPHSSTKVDKILPRGPTIYPC--PCKC-----PAPNLGSPSVF 124
QY 227 LFPKXPOTLMISKTPREVCVVVDVSHEDPEVKFNWYVDGVEVNAKTKPREEQYNSTYR 286
DB 125 IFPPKIDVIMISLSPIVTCVVVDVSEDDPDVQSWFNVNVEVHTAQOTHRBEDYNSTLR 184
QY 287 VVSIVLTALHDDMLNGKVKCKVSNKALPAPIEKTIKSAKQPRPQVYTLPPSRDELTKN 346
DB 185 VVSALPIQHDWMSGKFKCKVNNKDLPAPIERTISPKKSVAPQVYVLPPEEEMTKK 244
QY 347 QVSLTCLVKGFYPSDIAVENESNQPENNYKTTTPVLDSDGSFYLKSLTVDKSRWQGN 406
DB 245 QVTLTCNVTPMEDIVENTNNGKTELANKTEBPVLDSDGSYMYKLRVEKNWERN 304
QY 407 VFGSCVNHKALHNHYTOKSLSLSPGLQLDCTCAEADQDELGLWTT 452
DB 305 SYSQSVVHEGLHNHHTTKSPSRTPGGLDDVCAEAQGEGLDGLWTT 350

RESULT 15
I47159
Ig gamma 2a chain constant region - pig (fragment)
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C/Accession: I47159
R/Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994

QY 221 G-GPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 279
 DB 63 GPGSAPFIFPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 122
 QY 280 QYNSTYRVVSVLTVLIHQDLNKEKKYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPS 339
 DB 123 QFNSTYRVVSVLTPIQHOMLNGKKEKFKCVNKKDLPAPIRTRISAKAGQTRFEPQVYTLPP 182
 QY 340 RDELTKNOVSLTCLVKGFPYPSDIAVEMESNGQ--PENNYKTTTPVLDSGSFELYSKLT 397
 DB 183 TEELSRKSVKTLCLVLTGTPPIDVEMQPNQGPPEEGNYRTTPPOQDVDTGYFLYSKLA 242
 QY 398 DKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPG 431
 DB 243 DKASWQGRDGTQCAVMHEALHNHYTQKSISKTPG 276

RESULT 18

147160
 Ig gamma 2b chain constant region - pig (fragment)
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
 C:Accession: 147160
 R:Kacskovics, I.; Sun, J.; Butler, J.E.
 J:Immunol. 153, 3565-3573, 1994
 A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
 A:Reference number: 147158; MUID:95015845; PMID:7930579
 A:Accession: 147160
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-328 <KAC>
 A:Cross-references: EMBL:U03780; NID:9433125; PIDN:AAA52218.1; PID:9433126
 C:Genetics:
 A:Gene: IgG2b
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 F:133-202/Domain: immunoglobulin homology <IMW>

Query Match 33.2%; Score 898; DB 2; Length 328;
 Best Local Similarity 62.8%; Pred. No. 2.6e-46;
 Matches 172; Conservative 33; Mismatches 43; Indels 26; Gaps 6;

QY 168 TLVSQLELDSDGTCTVLTQ--NOKKVEFKID-----IVCPAPEPKSCDKHTHCP 220
 DB 70 TVPSSL--SSKSYTCNV--NHPTATTKVDKRGVTKTKPCP-----ICPACE 113
 QY 221 G-GPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 279
 DB 114 SPGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 173
 QY 280 QYNSTYRVVSVLTVLIHQDLNKEKKYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPS 339
 DB 174 QFNSTYRVVSVLTPIQHOMLNGKKEKFKCVNKKDLPAPIRTRISAKAGQTRFEPQVYTLPP 233
 QY 340 RDELTKNOVSLTCLVKGFPYPSDIAVEMESNGQ--PENNYKTTTPVLDSGSFELYSKLT 397
 DB 234 AEELSRKSVKTLCLVLTGTPPIDVEMQPNQGPPEEGNYRTTPPOQDVDTGYFLYSKRSV 293
 QY 398 DKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPG 431
 DB 294 DKASWQGGIFQCAVMHEALHNHYTQKSISKTPG 327

RESULT 19

19 gamma-2 chain C region - guinea pig
 C:Species: Cavia porcellus (guinea pig)
 C>Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 16-Jul-1999
 C:Accession: A94553; A90352; A90359; A90384; A90385; A02151
 R:Trischmann, T.M.
 Submitted to the Acllas, April 1975
 A:Reference number: A94553
 A:Accession: A94553

A:Molecule type: protein
 A:Residues: 1-3 <TRI>
 R:Birshlein, B.K.; Hussain, Q.Z.; Cebra, J.J.
 Biochemistry 10, 18-25, 1971
 A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. Am
 A:Reference number: A90352; MUID:71058471; PMID:5538606
 A:Accession: A90352
 A:Molecule type: protein
 A:Residues: 4-68 <BIR>
 R:Turner, K.J.; Cebra, J.J.
 Biochemistry 10, 9-17, 1971
 A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). II. Am
 A:Reference number: A90359; MUID:71058486; PMID:5538616
 A:Accession: A90359
 A:Molecule type: protein
 A:Residues: 69-133,312-329 <TUR>
 R:Tracey, D.E.; Cebra, J.J.
 Biochemistry 13, 4796-4803, 1974
 A:Title: Primary structure of the C-H2 homology region from guinea pig IgG2 antibodies.
 A:Reference number: A90384; MUID:75036073; PMID:4429665
 A:Accession: A90384
 A:Molecule type: protein
 A:Residues: 134-226 <TRA>
 R:Trischmann, T.M.; Cebra, J.J.
 Biochemistry 13, 4804-4811, 1974
 A:Title: Primary structure of the C-H3 homology region from guinea pig IgG2 antibodies.
 A:Reference number: A90385; MUID:75036073; PMID:4609467
 A:Accession: A90385
 A:Molecule type: protein
 A:Residues: 227-311 <TR2>
 R:Oliveira, B.; Lamm, M.E.
 Biochemistry 10, 26-31, 1971
 A:Title: Interchain disulfide bridges of guinea pig gamma-2- immunoglobulin.
 A:Reference number: A90354; MUID:71058474; PMID:4922544
 A:Contents: annotation; disulfide bonds
 A>Note: Cys-16 is involved in a heavy-light chain bond
 A>Note: Cys-105, Cys-107, and Cys-110 form inter-heavy chain bonds
 C:Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs.
 C:Complex: An immunoglobulin heterotrimeric subunit consists of two identical light (Lp)
 hain disulfide bonds. In some cases, such as IgG and IgM, the subunits associate into la
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin
 F:21-81/Domain: immunoglobulin homology <IM1>
 F:135-204/Domain: immunoglobulin homology <IM2>
 F:241-310/Domain: immunoglobulin homology <IM3>
 F:28-79/Disulfide bonds: #status experimental
 F:142-202/Disulfide bonds: #status experimental
 F:178/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:248-308/Disulfide bonds: #status experimental

Query Match 33.1%; Score 894.5; DB 1; Length 329;
 Best Local Similarity 62.5%; Pred. No. 4.2e-46;
 Matches 172; Conservative 34; Mismatches 60; Indels 9; Gaps 4;

QY 163 IOGKTLVSQLELDSDGTCTVLTQ--NOKKVEFKIDIVCPAPEPKSCDKHTHCP 218
 DB 58 IQSLGLVSLTSMVTPSSQKATCVVAHPASTKVDKTEPIRTPEPBCOTCK--CPPE 114
 QY 219 LIGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 278
 DB 115 NLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 174
 QY 279 QYNSTYRVVSVLTVLIHQDLNKEKKYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPP 338
 DB 175 QFNSTYRVVSVLTPIQHOMLNGKKEKFKCVNKKDLPAPIRTRISAKAGQTRFEPQVYTLPP 234
 QY 339 SRDELTKNOVSLTCLVKGFPYPSDIAVEMESNGQ--PENNYKTTTPVLDSGSFELYSKLT 396
 DB 235 SRDELTKNOVSLTCLVKGFPYPSDIAVEMESNGQ--PENNYKTTTPVLDSGSFELYSKLT 294
 QY 397 VKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPG 431
 DB 295 VDKSAMQGTIVTTCVSNHEALHNHYTQKSISKTPG 329

RESULT 20

Ig gamma 1 chain constant region - pig (fragment)
 147158
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
 C:Accession: I47158
 R:Kacskovics, I.; Sun, J.; Butler, J.E.
 J. Immunol. 153, 3565-3573, 1994
 A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
 A:Reference number: I47158; MUID:95015845; PMID:7930579
 A:Accession: I47158
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-328 <KAC>
 A:Cross-references: EMBL:U03778; NID:9433121; PIDN:AAA52216.1; PID:9433122
 C:Gene: IGCI
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 32.5%; Score 879.5; DB 2; Length 328;
 Best Local Similarity 55.2%; Pred. No. 3.3e-45;
 Matches 180; Conservative 38; Mismatches 67; Indels 41; Gaps 7;

QY 147 PGSSPVGCRSPRGKNIQGGKT-----LSVQLELDQSGTWC-----TVLQNK 191
 DB 2 PRTAPSVYPLPCGRVDSGPVALGCLASYPPEPVLTWNSGALTSVHTPPSVLQPSG 61
 QY 192 KYEFKIDI-VPCPAPEPKS--CDKTH-----TCPELG-----GPSVL 227
 DB 62 LVSLSSMTVVPASSLSKSYTCNVNHPATTKVDKRVGHIHQPTGICGCGEAVGSAVFI 121
 QY 228 PPPPKQDTLMISRTPEVTCVVDVSHEDPEVKFNMYDGEVNAATKPREQYNSTRV 287
 DB 122 PPPPKQDTLMISQTPETTCVVDVSKHAEVQFSWYDGEVHTAETRPKEQFNSTRV 181
 QY 288 VSVLTVLHODMUNGKRYKCVSKALPAPLEKTISSAKQOPREPOVYTLPPSDELTKQV 347
 DB 182 VSVLPTQHDQWLGKFKCKVNNVDLPAPLTRISAIQSRPQVYTLPPPAEELSRK 241
 QY 348 VSLTCLVKGFPYSDIAVESNSQ--PENNYKTPPVLDSDGFFLYSKLTVDKSRMOQG 405
 DB 242 VTLTCLVIGFPYDINVEKSNQOPREPTVTRTPPOVDGTFLYSKLAADKARNMDG 301
 QY 406 NVFSCVMEALHNHYTKSLSPG 431
 DB 302 DKFECAVMEALHNHYTKSLSPG 327

RESULT 21

IG gamma-2b chain C region, membrane-bound form - mouse
 G2MSBM
 C:Species: Mus musculus (house mouse)
 C>Date: 17-Dec-1982 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C:Accession: C02154; A02158; B02157
 R:Yamawaki-Kataoka, Y.; Nakai, S.; Miyata, T.; Honjo, T.
 Proc. Natl. Acad. Sci. U.S.A. 79, 2623-2627, 1982
 A:Title: Nucleotide sequences of gene segments encoding membrane domains of immunoglobul
 A:Reference number: A02154; MUID:82222190; PMID:6283537
 A:Accession: C02154
 A:Molecule type: DNA
 A:Residues: 335-405 <YAM>
 A:Cross-references: GB:J00462
 R:Rogers, J.; Choi, E.; Souza, L.; Carter, C.; Word, C.; Kuehl, M.; Eisenberg, D.; Wall,
 Cell 26, 19-27, 1981
 A:Title: Gene segments encoding transmembrane carboxyl termini of immunoglobulin gamma C
 A:Reference number: A02158; MUID:82115295; PMID:6799207
 A:Accession: A02158
 A:Molecule type: DNA
 A:Residues: 335-378 <ROG>
 A:Note: the translation of the first exon of the membrane-bound segment is given

R:Yamawaki-Kataoka, Y.; Kataoka, T.; Takahashi, N.; Obata, M.; Honjo, T.
 Nature 283, 786-789, 1980
 A:Title: Complete nucleotide sequence of immunoglobulin gamma2b chain gene cloned from r
 A:Reference number: A02157; MUID:80120716; PMID:6766534

A:Comment: a allele
 A:Accession: B02157
 A:Molecule type: DNA
 A:Residues: 1-335, 'K' <YAZ>

C:Comment: The sequence of residues 1-334 was assumed to be identical with the correspo
 C:Comment: Cell lines producing IgG contain two mRNA species for Ig gamma chains. The me
 hac it contains an alternative 3' end, encoded in separate exons, that is homologous wit
 C:Genetics:

A:introns: 1/1; 98/1; 120/1; 230/1; 335/1; 378/3

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kai
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1e
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobl
 F:143-212/Domain: immunoglobulin homology <IMM>
 F:352-369/Domain: transmembrane #status predicted <TM>
 F:370-405/Domain: intracellular #status predicted <INT>
 F:186/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 32.5%; Score 877.5; DB 1; Length 405;
 Best Local Similarity 51.7%; Pred. No. 5.5e-45;
 Matches 178; Conservative 52; Mismatches 75; Indels 39; Gaps 9;

QY 133 LIQG---QSLTTLSPSS-----PSVQCRSPRGKNIQGGKTVSGLQDQSGTW-- 182
 DB 28 LVKGYFPSSVYTWNSGLSSSVHTFPAL-----LQSG-LYTMSSTVTPSPSTWPS 77
 QY 183 ---TCTVLQ--NQKVEFKID-----IVPCAPPEKSCDTHTC--PELIGSPVFLF 228
 DB 78 QVTGVSVAHPASTVDKLEPSGPISTINPC-----PCKECHCPANLEGSPVFI 132
 QY 229 PPPPKQDTLMISRTPEVTCVVDVSHEDPEVKFNMYDGEVNAATKPREQYNSTRV 288
 DB 133 PENIKDVLIMISLTPKVTCCVVDVSEDDPVOQFSWYDGEVHTAETRPKEQFNSTRV 192
 QY 289 SVTLVHODMUNGKRYKCVSKALPAPLEKTISSAKQOPREPOVYTLPPSDELTKQV 348
 DB 193 STLPLOHDQWMSGKFKCKVNNKDLPSPLERTISKIGLVRAPOVYTLPPPAEQLSRKV 252
 QY 349 SLTCLVKGFPYSDIAVESNSQ--PENNYKTPPVLDSDGFFLYSKLTVDKSRMOQGNV 408
 DB 253 SLTCLVGNPFDIVERTSNQHTENYKDTAPVLDSDGYITTKLNMKTSKMEKTDSEF 312
 QY 409 SCNVMEALHNHYTKSLSPGLQDQTCABADGELDGLWTT 452
 DB 313 SCNVMEALHNHYTKSLSPGLQDQTCABADGELDGLWTT 356

RESULT 22

RMC274
 T-cell surface glycoprotein CD4 - chimpanzee
 N:Alternate names: T-cell surface antigen T4/Leu 3
 C:Species: Pan troglodytes (chimpanzee)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
 C:Accession: B32722; A46534
 R:Cammerlind, D.; Seed, B.
 Cell 60, 747-754, 1990
 A:Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi
 A:Reference number: A32722; MUID:90182664; PMID:2107024
 A:Accession: B32722
 A:Molecule type: mRNA
 A:Residues: 1-432 <CAM>
 A:Cross-references: GB:M31135
 R:Pomsgaard, A.; Hirsch, V.M.; Johnson, P.R.
 Eur. J. Immunol. 22, 2973-2981, 1992
 A:Title: Cloning and sequences of primate CD4 molecules: diversity of the cellular rece
 A:Reference number: A46534; MUID:93049640; PMID:1425921
 A:Accession: A46534
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA

A:Residues: 3-399 <F0M>
 A:Note: sequence extracted from NCBI backbone (NCBI:118332)
 C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; T-cell; transmembrane protein
 F:1-432/Product: T-cell surface glycoprotein CD4 #status predicted <EMT>
 F:1-371/Domain: extracellular #status predicted <EXT>
 F:9-86/Domain: immunoglobulin homology <IM1>
 F:111-161/Domain: immunoglobulin homology #status atypical <IM2>
 F:191-274/Domain: immunoglobulin homology <IM3>
 F:296-347/Domain: immunoglobulin homology <IM4>
 F:372-399/Domain: transmembrane #status predicted <TM>
 F:396-432/Domain: intracellular #status predicted <INT>
 F:16-84,130-159,303-345/Dissulfide bonds: #status predicted
 F:271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.4%; Score 875; DB 1; Length 432;
 Best Local Similarity 97.7%; Pred. No. 8,4e-45;
 Matches 170; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 27 KVLGKGDVVELTCTASOKSIOFHWKSNQIKILGQSGFLTKGPKLMDRADSRSL 86
 DB 2 KVLGKGDVVELTCTASOKSIOFHWKSNQIKILGQSGFLTKGPKLMDRADSRSL 61

QY 87 WDOGNPFLIKNLKIEDSDYICEVEDQKEEYQVLVFGITANSSTHLQSGSLTLTLESP 146
 DB 62 WDOGNPFLIKNLKIEDSDYICEVEDQKEEYQVLVFGITANSSTHLQSGSLTLTLESP 121

QY 147 PGSSSVQCRSPRGKNIQSGKTLVSQLELDSGWTCTVLOQKKEFKIDIV 200
 DB 122 PGSSSVQCRSPRGKNIQSGKTLVSQLELDSGWTCTVLOQKKEFKIDIV 175

RESULT 23
 147161

Ig gamma 3 chain constant region - pig (fragment)
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
 C:Accession: I47161
 R:Kacskovics, I., Sun, J.; Butler, J.E.
 J. Immunol. 153, 3565-3573, 1994
 A>Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
 A:Reference number: I47158; MUID:95015845; PMID:7930579
 A:Accession: I47161
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-328 <KAC>
 A:Cross-References: EMBL:U03781; NID:9433127; PIDN:AAA52219.1; PID:9433128
 C:Genetics:
 A:Gene: IG33
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 F:133-202/Domain: immunoglobulin homology <IM>

Query Match 32.3%; Score 873; DB 2; Length 328;
 Best Local Similarity 57.6%; Pred. No. 8e-45;
 Matches 179; Conservative 37; Mismatches 61; Indels 34; Gaps 9;

QY 137 QSLTLT-----LESPPSSPVQCSPPG-KNIQSGKTLVSQLELDSGWTCTVLOQ 190
 DB 35 EPTVMTNMGALTSQVHTFPV--LQPSGLVSLSMVTPASSL---SSKSYTCV--NH 87

QY 191 KKEVERKID-----IVCPAPEPKSCDKHTHCPCL-LGSPSVFLPPPKDMLMISRT 242
 DB 88 PATTTKVDKRVGKTTPKPCP-----ICPCPEVAGSVVTFPPKPDMLMISRT 136

QY 243 EYTCVVAVDSHEDEPVKFNWYDGYEVNNAKTPREEOYNSTYRVVSVLYLHODWLNKG 302
 DB 137 EYTCVVAVDSHEDEPVKFNWYDGYEVNNAKTPREEOYNSTYRVVSVLYLHODWLNKG 196

QY 303 EYKCKVSKALPAPLEKTIKAKGPRRPQVYTLPPSDELTKNOVSLTGVKGPYEDI 362
 DB 197 EFKCKVNVNDLPAPITRTISKAIQGRBPQVYTLPPPEELSRSKVTVTCLVIGYPPDI 256

QY 363 AVEWESNGQ--PENNYKTPPVLDSDGSFFLYSKLTVDKSRMOQGNVFGSVMEHALNH 420
 DB 257 HVEKSKSGQEPENNYKTPPVLDSDGSFFLYSKLTVDKSRMOQGNVFGSVMEHALNH 316

QY 421 YTKQSLSLSPG 431
 DB 317 YTKQSLSLSPG 327

RESULT 24
 G3MNC

Ig gamma-3 chain C region, secreted form - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Mar-1987 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
 C:Accession: B02156
 R:Wels, J.A.; Word, C.J.; Rimm, D.; Der-Balan, G.P.; Martinez, H.M.; Tucker, P.W.; Blatch
 EMBO J. 3, 2041-2046, 1984
 A>Title: Structure analysis of the murine IgG3 constant region gene.
 A:Reference number: A02156; MUID:85027161; PMID:6092053
 A:Accession: B02156
 A:Molecule type: DNA
 A:Residues: 1-329 <MEL>
 A:Cross-References: GB:U00451
 A:Note: the sequence was determined from the germline gene
 C:Genetics:
 A:Insertions: 97/1; 113/1; 223/1
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa)
 chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 F:19-83/Domain: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
 F:97-112/Region: hinge
 F:136-205/Domain: immunoglobulin homology <IM2>
 F:242-309/Domain: immunoglobulin homology <IM3>
 F:179,322/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.7%; Score 857.5; DB 1; Length 329;
 Best Local Similarity 59.6%; Pred. No. 6,7e-44;
 Matches 162; Conservative 38; Mismatches 57; Indels 15; Gaps 4;

QY 170 SVSGLQLDSDGTW----TCTVLOQKKEFKIDI--VCPAPEPKSCDKHTHCP--EL 219
 DB 62 SLSELVAVPSSSTWPSQTVICNVAPASKTELIRIKRIPKSPSPGS-----SCPGENI 116

QY 220 LGSPSVFLPPPKPDMLMISRTPEVTCVVVDVSHEDPEVFNNVVDGEVHNNAKTPREE 279
 DB 117 LGSPSVFLPPPKPDMLMISRTPEVTCVVVDVSHEDPEVFNNVVDGEVHNNAKTPREE 176

QY 280 QYNSTYRVVSVLYLHODWLNKGKCKVSKNKAAPLEKTIKAKGQPREPQVYTLPPS 339
 DB 177 QYNSTYRVVSVLYLHODWLNKGKCKVSKNKAAPLEKTIKAKGQPREPQVYTLPPS 236

QY 340 RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDK 399
 DB 237 RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDK 296

QY 400 SRMQQGNVFGSVMEHALNHNYTKSLSLSPG 431
 DB 297 DSWLQGEIFTCVSVVHEALNHNYTKSLSLSPG 328

RESULT 25
 PC4436

monoclonal antibody 13-1 heavy chain - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jan-2000
 C:Accession: PC4436
 R:Akashi, S.; Kato, K.; Torizama, T.; Dohmae, N.; Yamaguchi, H.; Kamachi, M.; Hazada, A.
 Biochem. Biophys. Res. Commun. 240, 566-572, 1997
 A>Title: Structural characterization of mouse monoclonal antibody 13-1 against a porphyr
 A:Reference number: JCS810; MUID:98063277; PMID:9398605
 A:Accession: PC4436
 A:Molecule type: protein

A:Residues: 1-444 <AKA>
 C/Comment: This catalytic antibody has peroxidase oxidase activity. It is directed again
 C/Superfamily: Immunoglobulin C region; immunoglobulin homology
 F:251-320/Domain: immunoglobulin homology <IMM>
 F:22/Denulfide bonds: Interchain (to 98) #status predicted
 F:99/Denulfide bonds: Interchain (to 109) #status predicted

Query Match 31.6%; Score 854.5; DB 2; Length 444;
 Best Local Similarity 41.5%; Pred. No. 1.4e-43;
 Matches 188; Conservative 76; Mismatches 118; Indels 71; Gaps 14;

QY 30 LGKGDVTELTCTAS--QKKSIOFHW-----KNSQIKL-----GNQGSFLTKGPS 74
 DB 11 LVRPGNSLKLSCTLSGFTFSNYRHMILRCPGKLEMIATVTKSDNYGAKVLESVRGR 70
 QY 75 KUNDRADSRSLWDQGNFPLIKNLKIEDSDTYIC-----EVED 113
 DB 71 TIS-RDSSKSSVYLQNM-----RLREEDPATYCCRTPMVYAMDCWQGTIVASSAKT 123
 QY 114 QKEEVOLLVFGLTANSDTHLQ-----QSLTTLTLESPGSSPSVQCRSPRKNIO 165
 DB 124 TPSPVYPLARGSAQNSMTLGLCKVGYPEPVYTNMS--GSLSSGVHTFPA---VLD 178
 QY 166 GKTLVSQLELDQSGTW-----TCTVLO--NQRKVEFKIDIVCPAPEPKSCDKHTCPE 218
 DB 179 SDLYTLSSSVTVPSSTWPSSETVTCNVAAHPASTKVDKIVRDCGC-KPCIC---TVE 233
 QY 219 LLAGPSVFLFPPPKKDTLMISRTPEVTCVVDVSHEDPEKFMWYVDGVEVNAKTRPRE 278
 DB 234 V---SSVFIFPPPKVLTITLTPKVCVVDISKDPEVQFSWFDVDEVHAQTPRE 290
 QY 279 EGYNSTYRVSVVTLVHODMLNGEKYCKVSNKALPAPIEKTISKAKGPREPOVYTLRP 338
 DB 291 EOPNSTFRSSELPPIHQDMLNGEKRCRNNSAPAPAPIEKTISKAKGKAPQVTTTPR 350
 QY 339 SRDELTKNOVSLTCLVKGFPSPDIAYWESNGOPENNYYKTPPEVLDSGSPFLYSKLTYD 398
 DB 351 PKQMAKDKVSLTCMTIDFPEDITVEMQNGOPAEYKNYKQITQIDTDSYFYFSLQNV 410
 QY 399 KSRWQGNVPSGCVWHEALNHTYOKSLSPG 431
 DB 411 KSNWEAGNTFTCSVLHEGLNHTKSLSPG 443

RESULT 26

S31459
 I9 gamma-1 chain - sheep (fragment)
 C/Spectrum: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
 C/Accession: S31459
 R:Parti, S.; Nau, F.
 Submitted to the EMBL Data Library, December 1992
 A/Reference number: S31459
 A/Accession: S31459
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-472 <PAT>
 A/Cross-references: EMBL:X69797
 C/Superfamily: Immunoglobulin C region; immunoglobulin homology
 C/Keywords: immunoglobulin
 F:277-346/Domain: immunoglobulin homology <IMM>

Query Match 31.6%; Score 854.5; DB 2; Length 472;
 Best Local Similarity 42.2%; Pred. No. 1.5e-43;
 Matches 196; Conservative 55; Mismatches 119; Indels 94; Gaps 16;

QY 36 TVELTCTAS--QKKSIOFHWKN---SNQIKILNGO--SFLTKGP---SKUNDRADSR 85
 DB 34 TISVLTCTISGSLANNVGVMDVVRQAPGKALEWLGSGYDEIDYNPLKSLITTKDTSS 93
 QY 86 LWDQGNFPLIKNLKIEDSDTYICEVEDQEE-----VOLLVFGLTANSDTHL 133
 DB 94 -----QVSLTLSTVTTEDTRAVVYCARVDYDSSHAFAVASYDFFWGPGLLISVLSASG----- 144

QY 134 LQGSUTLTLSPGSSPSVQCRSPRKNI-----QSGKTLVSQ 173
 DB 145 -----TPPKVYPLTSCGDPSSISVTLGLCVSSYMEPEVYTNMSALTSQVHT 193

QY 174 LE--LQDSG-----TWCTVLO--NQRKVEFKIDIVCPAPEPKSCD 211
 DB 194 FPAIQLSSGLYSLSSVTVTPASTSGAQTICNVAAHPASTKVDKIVE--DCGPD--CKHC- 250
 QY 212 KTHTC--PELLGSPVFLFPPPKKDTLMISRTPEVTCVVDVSHEDPEKFMWYVDGVE 269
 DB 251 ---RCPPPELPGPSVFIFPPPKKDTLTISGPEVTCVVDVQDDPEVQFSWFDVNEV 307
 QY 270 HNAKTRPEQYNSTYRVSVVTLVHODMLNGEKYCKVSNKALPAPIEKTISKAKGPR 329
 DB 308 FRARKTRPEQYNSTYRVSVVTLVHODMLNGEKYCKVSNKALPAPIEKTISKAKGPR 367
 QY 330 EPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYWESNGOP--ENNYKTPPEVLDSG 387
 DB 368 EPQVYTLAPQGEISKSLSTVCLVGFYPDVIAWQNGQPESEDKYGTTSOLDAG 427
 QY 388 SFFLYSKLTYDYSRWQGNVPSGCVWHEALNHTYOKSLSPG 431
 DB 428 SYFLSRLRLVDKNWSQEGDTVACVWHEALNHTYOKSLSPG 471

RESULT 27

S37483
 I9 gamma-2a chain - mouse
 C/Spectrum: Mus musculus (house mouse)
 C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
 C/Accession: S37483
 R:Ducancel, F.F.D.
 Submitted to the EMBL Data Library, February 1993
 A/Reference number: S37483
 A/Accession: S37483
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-469 <DUC>
 A/Cross-references: EMBL:X70423; NID:g406252; PIDN:CAA9868.1; PID:g406253
 C/Superfamily: Immunoglobulin C region; immunoglobulin homology
 C/Keywords: immunoglobulin
 F:276-345/Domain: immunoglobulin homology <IMM>

Query Match 31.5%; Score 850.5; DB 2; Length 469;
 Best Local Similarity 42.5%; Pred. No. 2.7e-43;
 Matches 197; Conservative 56; Mismatches 125; Indels 85; Gaps 15;

QY 30 LGKGDVTELTCTASQKKSIOFHWKNSQIKILNGO-----SFLTKGPSKUNDRADSR 83
 DB 30 LVRPGASVKISCKASGYFTDY---INWVKQKPGQGLKIMWIYPASGNTYNNFQK 86
 QY 84 RSL---WDQGNFPLIKNLKIEDSDTYICEVEDQEEVQLVFGLTANSDTHLQ--GQS 138
 DB 87 ATLTYDTSSTAYMQLSLTSIEDTAVFC-----ARAGATAT---LDDWGQ 132
 QY 139 LTLTLESPGSSPSVQ-----CRSPRKNIOGQ-----KTL----- 169
 DB 133 TLTIVSSAKTTPASVYPLAPVCGDTGSSVTLGLCKVGFPEPVTLTNMSGSLSGVHTP 192
 QY 170 -SVSQLELO-----DSGTW-----TCTVLO--NQRKVEFKIDIVCPAPEPK 208
 DB 193 PAVLQSDLTLSSTVTVSTWPSQSLTCNVAAHPASTKVDKIVRDCGPTIKPCP--PC 249
 QY 209 SCDKTHTCPELLGSPVFLFPPPKKDTLMISRTPEVTCVVDVSHEDPEKFMWYVDGVE 268
 DB 250 KC-----PAPNLGGSVFIFFPKKIDVLMISLPIVTCVVDVSDDDPVQVSWFVNNV 305
 QY 269 VNAKTRPREQYNSTYRVSVVTLVHODMLNGEKYCKVSNKALPAPIEKTISKAKGPR 328
 DB 306 VHTAQTOTREBYNSTLRVVSALPIQHDPMWGEKFCVNNKDLPAPIERTISKPKGSV 365
 QY 329 REPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYWESNGOPENNYYKTPPEVLDSG 388

Db 366 RAPQYVLPPEPEEETKQVLTTCVTFMPEPDIYVENNKGTELANKTEPVLDSG 425
 QY 369 FFLYSLKLTVDKSRWQGVNFCVSHVHEALHNHYTKSLSPG 431
 Db 426 YFMYSKLREKKNWERNYSYSCVSHVHEGLHNHYTKSLSPG 468

RESULT 28

C30554
 Ig heavy chain C region - sheep (fragment)
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 03-Mar-1989 #sequence_revision 03-Mar-1989 #text_change 21-Jan-2000
 C:Accession: C30554
 R:Foley, R.C.; Beh, K.J.
 J:Immunol. 142, 708-711, 1989
 A:Title: Isolation and sequence of sheep Ig H and L chain cDNA.
 A:Reference number: A30554; MUID:89093962; PMID:2492052
 A:Accession: C30554
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-308 <FOL>
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: heterotrimer; immunoglobulin
 F:113-182/Domain: immunoglobulin homology <IMM>

Query Match 31.0%; Score 838; DB 2; Length 308;
 Best Local Similarity 60.7%; Pred. No. 8.9e-43;
 Matches 164; Conservative 37; Mismatches 55; Indels 14; Gaps 7;

QY 170 SVSLELDQSG--TWCTVLQ--NOKYEFKIDIVCPAPEKSCDKHTC--PELLGAP 223
 Db 44 SVTVPASTGAGQTFICVNAHPASTKVDKRYE--GCGDP--CKHC----RCPPPELPGP 97
 QY 224 SVFLFPPPKKDTLMSRTPEVTCVVDVSHEDPEKFMVYDGVVNAKTPREEOYNS 283
 Db 98 SVFIFPPPKKDTLISGTPVTCVVDVQDDPEVQFPMFVNVNVEVTAETKPREEOYNS 157
 QY 284 TYRVSVLTVLHQMNGKEYCKVSNKALPAPIKTSISKAGQPREOYVTLPPSRDEL 343
 Db 158 TFRVVSALPIQHDMGKEFKCKVSNKALPAPIKTSISKAGQPREOYVTLPPSRDEL 217
 QY 344 TKNOVSLTCLVKGFYPSDIAVWESNGQP--ENNYKTTTPVLDSGSEFLYSLKLTVDKSR 401
 Db 218 SKSTLSVTCVLTGFPDYIAVEMQNGQDESEDKYGTTSQDADGSEFLYSRLKLVDSNS 277
 QY 402 WQGNVFCSCVHHEALHNHYTKSLSPG 431
 Db 278 WQEGDTYACVWHEALHNHYTKSLSPG 307

RESULT 29

PS0018
 Ig gamma-2b chain C region - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Jul-1999
 C:Accession: PS0018; B25941
 R:Brueggemann, M.
 Gene 74, 473-482, 1988
 A:Title: Evolution of the rat immunoglobulin gamma heavy-chain gene family.
 A:Reference number: PS0017; MUID:89232738; PMID:3149946
 A:Accession: PS0018
 A:Molecule type: DNA
 A:Residues: 1-333 <BRU>
 R:Brueggemann, M.; Free, J.; Diamond, A.; Howard, J.; Cobbold, S.; Waldmann, H.
 Proc. Natl. Acad. Sci. U.S.A. 85, 6075-6079, 1988
 A:Title: Immunoglobulin heavy chain locus of the rat: striking homology to mouse antibody
 A:Reference number: A25941; MUID:86287297; PMID:3016742
 A:Accession: B25941
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 227-333 <BR2>
 C:Genetics:

A:introns: 96/1; 117/1; 227/1
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: immunoglobulin
 F:20-82/Domain: immunoglobulin homology <IMM>

Query Match 30.8%; Score 833; DB 2; Length 333;
 Best Local Similarity 58.6%; Pred. No. 1.9e-42;
 Matches 157; Conservative 38; Mismatches 53; Indels 20; Gaps 5;

QY 179 SGTW-----TCTVQ--NOKYEFKID-----VPCPAPPEKSCDKHTC--PELLGAP 223
 Db 70 SFTVPSQTVTCVNAHPASTKVDKVERRNGGIGHKCP----TCPTCKCPVPELLGAP 124
 QY 224 SVFLFPPPKKDTLMSRTPEVTCVVDVSHEDPEKFMVYDGVVNAKTPREEOYNS 283
 Db 125 SVFIFPPPKKDTLISQNAKVTGVVDVSEEPVQFPMFVNVNVEVTAETKPREEOYNS 184
 QY 284 TYRVSVLTVLHQMNGKEYCKVSNKALPAPIKTSISKAGQPREOYVTLPPSRDEL 343
 Db 185 TFRVVSALPIQHDMGKEFKCKVSNKALPAPIKTSISKAGQPREOYVTLPPSRDEL 244
 QY 344 TKNOVSLTCLVKGFYPSDIAVWESNGQPRENNYKTTTPVLDSGSEFLYSLKLTVDKSRWQ 403
 Db 245 TEQTVSLTCLTSGFLPNDIGVETSNCHIEKNYKTEPVMDSDGSEFLYSRLKLVDSNS 304
 QY 404 QGNVFCSCVHHEALHNHYTKSLSPG 431
 Db 305 SRAVFCVSHVHEGLHNHYTKSLSPG 332

RESULT 30

S40295
 Ig gamma-2a chain (mAb735) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 16-Jul-1999
 C:Accession: S40295
 R:Klebert, S.; Kratzin, H.D.; Zimmermann, B.; Vaesen, M.; Frosch, M.; Weisgerber, C.; Bi
 submitted to the EMBL Data Library, January 1993
 A:Description: Primary structure of the murine monoclonal IgG2a antibody mAb735 against
 A:Reference number: S40295
 A:Accession: S40295
 A:Molecule type: protein
 A:Residues: 1-446 <KLE>
 C:Genetics:
 A:Map position: 12
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: disulfide bond; glycoprotein; immunoglobulin; pyroglutamic acid
 F:1-446/Product: Ig gamma-2a chain #status experimental <MAT>
 F:1-117/Domain: V-D-J region <VDJ>
 F:118-446/Domain: C region <CHR>
 F:118-214/Domain: C1 region <CH1>
 F:215-230/Region: hinge
 F:231-340/Domain: C2 region <CH2>
 F:341-446/Domain: C3 region <CH3>
 F:360-427/Domain: immunoglobulin homology <IMM>
 F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
 F:22-96, 144-199, 261-321, 367-428/Disulfide bonds: #status predicted
 F:113/Disulfide bonds: interchain (to light chain) #status predicted
 F:224, 227, 229/Disulfide bonds: interchain #status predicted
 F:297/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 30.5%; Score 825; DB 2; Length 446;
 Best Local Similarity 40.9%; Pred. No. 8.2e-42;
 Matches 190; Conservative 60; Mismatches 123; Indels 92; Gaps 15;

QY 30 LGKKGDTVELTCTASQKSIQF--HWKNSNQIKILNQG-----SFLYKGPSKLNDRAD 81
 Db 11 LVRRGASVSKSKSGVTFDYIHW-----VKRRPEGGELWIMGIVTSGNTKYNKFX 65
 QY 82 SRSL---WQGNPPLIKLKIEDSPYICEVVDQKEVGLVFGLTANSDTLLQ--G 136
 Db 66 GKATLTVDTSSTAYWQSLSTSEDSAVYFC-----ARGSKPAMDYWG 108

```

0Y 137 QSLTTLTSPGSSSVG-----CSRPGRXNIQGS-----KTL----- 169
    |:::|:::|:::|
Db 109 QGTSVTVSSAKTTABSVPLAPVCGDITGSSVTLGCLVKGYFPEPVTLTMNSSLGSSGVH 168
    |:::|:::|:::|
0Y 170 ---SVSOLELO-----DSGTW-----TCTVLQ--NOKRYEFKID-----IYCPAPE 206
    |:::|:::|:::|
Db 169 TFPVAVLGSDLTLTSSVTVTSSITWPSOSITCNVAHPASTKYDKKIERGRTIIFCP--- 225
    |:::|:::|:::|
0Y 207 PKSCDKHTHTCELLGSPSVFLFPKPDKDTLMISRPEVTCVVVDVSHEDPEVKENRYVDG 266
    |:::|:::|:::|
Db 226 PCKC----PARNLLGGSPSVFIIPPKIKDVLNMLSIPMTCCVVVDSEDDPVOGTSWIFNN 281
    |:::|:::|:::|
0Y 267 VEVNNAKTKRPEEOYNSTIRYVSULTIADQWLNGKEYKCKVSKALPAPIEKTISKAG 326
    |:::|:::|:::|
Db 282 VEVLTAQTQTHREDYNSLTRVVSALPIQHOMWSGKEFKCKVKNNKDLPAPIERTISKPG 341
    |:::|:::|:::|
0Y 327 QPREPOVYTLTPSRDELTKNOVSLTCLVKGYPYSIAIEMWSNQGPENNYKTTPEPVLSD 386
    |:::|:::|:::|
Db 342 SVRAPQVVLPPEPEEMTKQVTLTCVATDMPEDIIYEWNNNGTTELANTPEVLSD 401
    |:::|:::|:::|
0Y 387 GSFPLYSKLTVDKSRMOOGNIFSCVMHEALHNHYTOKSLSLSPG 431
    |:::|:::|:::|
Db 402 GSYFMYSKLTARVEKKNWERNYSYCSVHNEGHLNHHHTTKISFRTPG 446
    |:::|:::|:::|

```

```

RESULT 31
PS0017
Ig gamma-1 chain C region - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Jul-1999
C/Accession: PS0017; C25941
R/Bruggemann, M.
Gene 74, 473-482, 1988
A/Title: Evolution of the rat immunoglobulin gamma heavy-chain gene family.
A/Reference number: PS0017; MUID:89232738; PMID:3149946
A/Accession: PS0017
A/Molecule type: DNA
A/Residues: 1-326 <BRU>
R/Bruggemann, M.; Frey, J.; Diamond, A.; Howard, J.; Cobbold, S.; Waldmann, H.
Proc. Natl. Acad. Sci. U.S.A. 83, 6075-6079, 1986
A/Title: Immunoglobulin heavy chain locus of the rat: striking homology to mouse antibody
A/Reference number: A25941; MUID:86287397; PMID:3016742
A/Accession: C25941
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 220-326 <BR2>
C/Genetics:
A/Intons: 98/1; 113/1; 220/1
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: immunoglobulin
;20-84/Domain: immunoglobulin homology <IM>

```

[illegible]

	240	406	300
Db	NEWSIICMKGFEYPPDIYVWQMNOQPOENYKNTPTIMDTGSLFYLSKUNVKKERWQQG		
Qy		NVPSCSVMHEALHNHYTKQSLSPG	431
Db		NTFTCSVLHGBLHNHHTKSLSHSPG	325

RESULT 32
S00847
Ig gamma-2c chain C region - rat (fragment)
C:Species: Rattus norvegicus (Norway Rat)
C:Date: 01-Dec-1989 #sequence Revision 01-Dec-1989 #text change 23-Jul-1999
C:Accession: S00847
R:Bruggemann, M.; Delmaestro-Galfre, P.; Waldmann, H.; Calabi, F.
E:Eur. J. Immunol. 18, 317-319, 1988
A:Title: Sequence of a rat immunoglobulin gamma-2c heavy chain constant region cDNA: ex
A:Reference numbers: S00847; MUID:86166903; PMID:3127222
A:Accession: S00847
A:Molecule type: mRNA
A:Residues: 1-329 <BRU>
A:Cross-references: EMBL:X07189; NID:G57602; PIDD:CAA30169.1; PID:g663228
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:20-84/Domain: immunoglobulin homology <IMM>

Query	March 30.5%;	Score	823;	DB	2;	Length	329;
Best Local Similarity	58.5%;	Pred. No.	7.5e-42;				
Matches	151;	Conservative	45;	Mismatches	56;	Indels	6;
				Gaps	2;		
Qy	179	SGTW-----TCTVLGNOKKVEFKIDIVCPAPPEKSCDKHTPCDELLGSEVFLFPKPK	233				
Db	72	SSWSSQYTCSVAHPATKSNLIKIEP-RRPKRPRPDICSCDNLGRSRVFLFPKPK	130				
Qy	234	DTLMSIRTEPVTCVVVDVSHEDPEVKFMWYVDGVEVHNAKTPREEDYNSITYRVSVLTV	293				
Db	131	DILMITLTFRKTCVVVDVSEEDPVQFSPFMFVDNVRVFLAQIQPHEQLGTFPRVSLTHI	190				
Qy	294	LHODWINGEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKDVNSITCL	353				
Db	191	QHODMWSGEKFCCKVNNKDLPSPIEKTISKDRGKARTRPQVYTIIPPRRQSNKNVSLTCL	250				
Qy	354	VKGFYPSDLAVEMESNGDPENNYKTTTPVLDSDGSFLYSLTLTYDKSRMQQGNVFGSSVM	413				
Db	251	VTSFYPSASISVEMERNGELDEQYKNTLTVLDSDESFLYSLKLSDTDSMMKGDYITCSVV	310				
Qy	414	HEALTHNYTKQSLISPG	431				
Db	311	HEALTHHHYTKQLSRSPG	328				

RESULT 33
GIMS
Ig gamma-1 chain C region, secreted form - mouse
C/Species: Mus musculus (house mouse)
C/Date: 31-Mar-1980 #sequence revision 24-Sep-1981 #text_change 16-Jul-1999
C/Accession: A02159; A26234; A26236
R/Honjo, T.; Obata, M.; Yamawaki-Kataoka, Y.; Kataoka, T.; Kawakami, T.; Takahashi, N.;
Cell 18, 559-568, 1979
A/Title: Cloning and complete nucleotide sequence of mouse immunoglobulin gamma1 chain
A/Reference number: A02159; PMID:80045036; PMID:115593
A/Accession: A02159
A/Molecule type: DNA
A/Residues: 1-324 <HON>
A/Cross-references: GB:J00453
A/Note: The sequence was determined from the germ-line gene
A/Note: Lys-324 is removed posttranslationally
R/Obata, M.; Yamawaki-Kataoka, Y.; Takahashi, N.; Kataoka, T.; Shimizu, A.; Mano, Y.; Saito,
Gene 9, 87-97, 1980
A/Title: Immunoglobulin gamma-1 heavy chain gene: structural gene sequences cloned in a
A/Reference number: A26234; PMID:80202559; PMID:6769752
A/Contents: MOPC 31C
A/Accession: A26234

```

A:Molecule type: mRNA
A:Residues: 76-324 <OBA>
A:Cross-references: GB:V00775; NID:951652; PIDN:CAA24153.1; PID:951653
R:Rogers, J.; Clarke, P.; Salsner, W.
Nucleic Acids Res. 6, 3305-3321, 1979
A>Title: Sequence analysis of cloned cDNA encoding part of an immunoglobulin heavy chain
A:Reference number: A26236; MUID:80012837; PMID:113776
A:Contents: MOPC 21
A:Accession: A26236
A:Molecule type: mRNA
A:Residues: 170-275, 'D', 277, 'D', 279-322 <ROG>
A:Cross-references: GB:V00795; NID:951830; PIDN:CAA24176.1; PID:9780265
R:Adetugbo, K.
J. Biol. Chem. 253, 6068-6075, 1978
A>Title: Evolution of immunoglobulin subclasses. Primary structure of a murine myeloma g
A:Reference number: A26237; MUID:78242288; PMID:98524
A:Contents: annotation; MOPC 21
A>Note: This is the final paper in a series reporting the protein sequence, the disulfid
A>Note: There are a number of differences from the sequence shown
C:Genetics:
A:Introns: 1/1; 98/1; 111/1; 218/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into la
C:Superfamily: Immunoglobulin C region; Immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob
F:20-84/Domain: Immunoglobulin homology <IM1>
F:98-110/Region: hinge
F:131-200/Domain: Immunoglobulin homology <IM2>
F:237-304/Domain: Immunoglobulin homology <IM3>
F:27-82, 138-198, 244-302/Disulfide bonds: #status experimental
F:102/Disulfide bonds: Interchain (to light chain) #status experimental
F:104, 107, 109/Disulfide bonds: Interchain (to heavy chain) #status experimental
F:174/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query March 30.3%; Score 818.5; DB 1; Length 324;
Best Local Similarity 51.8%; Pred. No. 1.4e-41;
Matches 156; Conservative 50; Mismatches 60; Indels 35; Gaps 5;

Qy 133 LLQGSGLTLTLESPGSSPVQCRSPRGKNIQGGKTLISVQLQDSDGTMTCTVLQ--NQ 190
Db 56 VLQSDVLTLSSTSVTVVSSP-----RPSTVTVCNVAHPASS 90
Qy 191 KKEPEKIDIVCPAPEPEKSCDKHTTCPELLGSPVFLFPPPKDITMISRTPEVTCVVD 250
Db 91 TKVDKKTIVPRDCCG-KPCIC---TVPEV---SSVFIRPPKPKDVLITTLTPKTCVVD 142
Qy 251 VSHDEPEKFMVYDGVGVNNAKTRREOVNSTRVSVLTVLHOMDLNKEKCKKYSN 310
Db 143 ISKDPPEVQFSWPDVDVHTAQTQPREQNSSTRSVSELPIMHODMLNKEKCRVNS 202
Qy 311 KALPAIEKTSKXAGQPREPQVYTLTPSRDELTKNQVSLTCLVKGFYPSDIAVMEESNG 370
Db 203 AAFPAPIEKTSKRGKAPQVYTIIPPKQMAKDKSLCTMTDPEPDITVEMQNG 262
Qy 371 QPENNYKTPVLDSDGSFFLYSKLTVDKSRWQGNVFCGVMEHALNHYTQSLISLP 430
Db 263 QPAENYKATQIPIMNTNGSYFYYSKLVNQKSNMEAGNFTCSVLHGLAHNHTKSLSHSP 322
Qy 431 G 431
Db 323 G 323

RESULT 34
G2M511
Ig gamma-2b chain - mouse
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1990 #sequence revision 01-Dec-2000 #text change 01-Dec-2000
C:Accession: S25057; A02157; A26235; A26232; A26233; A53598
R:Flischer, R.; Voss, A.; Niersbach, M.; Munkler, W.; Hirsch, H.J.; Kreuzaler, F.
submitted to the EMBL Data Library, July 1992
A:Description: Production of a Tobacco mosaic virus (TMV) inactivating neotop specific m
A:Reference number: S25057

```

```

A:Accession: S25057
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-474 <FTS>
A:Cross-references: EMBL:X67210; NID:954826; PIDN:CAA47649.1; PID:954827
R:Yanawaki, K.; Katoh, Y.; Katoh, T.; Takahashi, N.; Obata, M.; Honjo, T.
Nature 283, 786-789, 1980
A>Title: Complete nucleotide sequence of immunoglobulin gamma2b chain gene cloned from n
A:Reference number: A02157; MUID:80120716; PMID:6766534
A:Contents: a allele
A:Accession: A02157
A:Molecule type: DNA
A:Residues: 138-161, 'L', 163-189, 'FP', 193-474 <YAM>
A:Cross-references: GB:J00461
A>Note: The sequence was determined from the germ-line gene
R:Tucker, P.W.; Marcu, K.B.; Slightom, J.L.; Blattner, F.R.
Science 206, 1299-1303, 1979
A>Title: Structure of the constant and 3' untranslated regions of the murine gamma2b hea
A:Reference number: A26235; MUID:80081501; PMID:117548
A:Contents: MPC 11
A:Accession: A26235
A:Molecule type: mRNA
A:Residues: 138-172, 'P', 174-189, 'FP', 193-376, 'T', 378-474 <TU1>
A>Note: Lys-474 is probably removed posttranslationally
R:Tucker, P.W.; Marcu, K.B.; Newell, N.; Richards, J.; Blattner, F.R.
Science 206, 1303-1306, 1979
A>Title: Sequence of the cloned gene for the constant region of murine gamma2b immunoglob
A:Reference number: A26232; MUID:80081502; PMID:117549
A:Accession: A26232
A:Molecule type: DNA
A:Residues: 138-172, 'P', 174-189, 'FP', 193-376, 'T', 378-474 <TU2>
R:Ollo, R.; Rougeon, F.
Nature 296, 761-763, 1982
A>Title: Mouse immunoglobulin allotypes: post-duplication divergence of gamma2a and gamm
A:Reference number: A26233; MUID:82179203; PMID:6803173
A:Contents: b allele
A:Accession: A26233
A:Molecule type: DNA
A:Residues: 138-161, 'L', 163-189, 'FP', 193-300, 'R', 302-331, 'A', 333-437, 'DI', 440-474 <OL1>
R:Kim, H.; Yamaguchi, Y.; Masuda, K.; Matsunaga, C.; Yamamoto, K.; Irimura, T.; Takahash
J. Biol. Chem. 269, 12345-12350, 1994
A>Title: O-glycosylation in hinge region of mouse immunoglobulin G2b.
A:Reference number: A53598; MUID:94216359; PMID:7512967
A:Accession: A53598
A:Status: preliminary
A:Molecule type: protein
A:Residues: 234-251 <KIM>
C:Comment: The a allele sequence is shown.
C:Genetics:
A:Introns: 138/1; 236/1; 258/1; 368/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into la
C:Superfamily: Immunoglobulin C region; Immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob
F:157-222/Domain: Immunoglobulin homology <IM1>
F:236-257/Region: hinge
F:281-350/Domain: Immunoglobulin homology <IM2>
F:387-454/Domain: Immunoglobulin homology <IM3>
F:152/Disulfide bonds: Interchain (to light chain) #status predicted
F:164-220, 288-348, 394-452/Disulfide bonds: #status predicted
F:247, 250, 253, 256/Disulfide bonds: Interchain (to heavy chain) #status predicted
F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query March 30.3%; Score 818; DB 1; Length 474;
Best Local Similarity 39.7%; Pred. No. 2.3e-41;
Matches 184; Conservative 66; Mismatches 124; Indels 90; Gaps 14;

Qy 34 GDTVELCTASQKSIQF--HWKNSNIOKITLNGG----SFL--TKGSPKNDPADSRSS 85
Db 34 GASYKNSCKASGTYFTYVNMW-----VKQKPGSGLEWIGYINNKGTGKFNKFKGKAT 88
Qy 86 LW--DQGNPPLIIKNLKIEDSTYICEVBDQKEEVQLVFGLTANSDTHLLQGSGLTLT 142

```

```

Db      89 LPSDKSNYAYMELSLTSDSAVYYCARN-----YDYMFAYWGGGLVT 134
Qy      143 LSPSPSSSVQ-----CSPGKNIQGG-----KTLVSQLE 175
Db      135 VSAAKTTPSVVPLAPGCDTGTSSVSCGLVKGYPPESVTVTNSGSLSSVHTLSQAL 194
Qy      176 LQD-----SGTW-----TCTVAG--NOKKVEKID-----IYCPAPEPS 209
Db      195 LQSGLYTSSSVTVSSSTPSCQTVTCVNHAPSSTTVDDKLEPSPGISTINPCP-----P 249
Qy      210 CDKTHTC--PELLGSPVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNYYDGV 267
Db      250 CKECHKCAPNLEGGSVIFPPNKKDVLMISLTPKTCVVVDVSEDDPDVQISMFVNNV 309
Qy      268 EVHNAKTKREEQYNSTVSVVLTVLHODMLNGEKYCKCKVSNKALPAPIETKISKAKQ 327
Db      310 EVHTAQOTQTHREDYNSTIRVSTLPIQHODMWSGKEFKCKVNNKDLPSPIERTISKIKGL 369
Qy      328 PREPOVYTLPPSRDELTKNVSILCLVKGFPSPDIAYVESNGOPENNYKTPVLDSDG 387
Db      370 VPAPOVYTLPPAEQUSKDVSLCLVGFNPGDISVETNSGHTENYKDTAPVLDSDG 429
Qy      388 SFFLYSKLTVDSRMQGNVFCSSVMEALHNNYTKSLSLSPG 431
Db      430 SFYISKLNMKTSKMEKTDTSFSCNVRHEGLKNYLLKKTISRSPG 473

```

RESULT 35

GMSA

```

Ig gamma-2a chain C region, secreted form (allele a) - mouse
C.Species: Mus musculus (house mouse)
C.Date: 30-Sep-1980 #sequence_revision 01-Sep-1981 #text_change 16-Jun-2000
C.Accession: A02152; A32657; A32658
R.Sikora, J.L.; Aufreay, C.; Rougeon, F.
Nucleic Acids Res. 8, 3143-3155, 1980
A.Title: Structure of the constant and 3' untranslated regions of the murine Balb/C gamma
A.Reference number: A02152; MUID:81076554; PMID:677755
A.Accession: A02152
A.Molecule type: mRNA
A.Residues: 1-330 <SIX>
A.Cross-references: GB:V00798; NID:G51835; PIDN:CNA24178.1; PID:G1333984
R.Ollo, R.; Aufreay, C.; Morchamps, C.; Rougeon, F.
Proc. Natl. Acad. Sci. U.S.A. 78, 2442-2446, 1981
A.Title: The complete nucleotide sequence of mouse immunoglobulin gamma 2 A gene and cDNA
A.Reference number: A32657; MUID:81198976; PMID:6262729
A.Accession: A32657
A.Molecule type: DNA
A.Residues: 1-330 <YAM>
A.Cross-references: GB:J00470
A.Note: The sequence was determined from the germ-line gene
R.Ollo, R.; Aufreay, C.; Morchamps, C.; Rougeon, F.
Proc. Natl. Acad. Sci. U.S.A. 78, 2442-2446, 1981
A.Title: Comparison of mouse immunoglobulin gamma-2A and gamma-2B chain genes suggests
A.Reference number: A32658; MUID:81233894; PMID:6787604
A.Accession: A32658
A.Molecule type: DNA
A.Residues: 1-330 <OLU>
A.Note: The sequence was determined from the germ-line gene
R.Bourgeois, A.; Fougereau, M.; Rocca-Serra, J.
Eur. J. Biochem. 43, 423-435, 1974
A.Title: Determination of the primary structure of a mouse IgG2a immunoglobulin: amino-
A.Reference number: A32659; MUID:7417517; PMID:4831970
A.Contents: annotation; myeloma protein MOPC 173
A.Note: this is one paper in a series reporting the sequence; for additional references,
Ride Preval, C.; Fougereau, M.
Eur. J. Biochem. 30, 452-462, 1972
A.Title: Determination of the primary structure of a mouse gamma2a immunoglobulin. Iden
A.Reference number: A32660; MUID:73056887; PMID:4565406
A.Contents: annotation; MOPC 173, disulfide bonds
C.Genetics:

```

```

A.Introns: 1/1; 98/1; 114/1; 224/1
C.Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C.Superfamily: immunoglobulin C region; immunoglobulin homology
C.Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
F:20-84/Domain: immunoglobulin homology <IM1>
F:98-113/Region: hinge
F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:15/Disulfide bonds: interchain (to light chain) #status experimental
F:27-82,144-204,250-308/Disulfide bonds: #status experimental
F:107,110,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Aen) (covalent) #status predicted

```

Query Match

```

29.8%; Score 805.5; DB 1; Length 330;
Best Local Similarity 58.5%; Pred. No. 8,26-41;
Matches 155; Conservative 33; Mismatches 58; Indels 19; Gaps 5;

```

```

Qy      179 SGTW-----TCTVLD--NOKKVEKID-----IYCPAPEPSCDKTHTCPELLGSPSVF 226
Db      72 SGTWSPQSITCVNAPASTKVKDKIIEBGRPIKCP---PCKC---PAPVLGGPSVF 124
Qy      227 LPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNYYDGVVHNAKTKPREEQYNSTYR 286
Db      125 IFPPKIKDVLMLSLSPITVTCVVVDVSEDDPDVQISMFVNNVEVHTAQOTQTHREDYNSTLR 184
Qy      287 VVSUTVLHODMLNGEKYCKCKVSNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKN 346
Db      185 VVSALPIQHODMWSGKEFKCKVNNKDLPAPIERTISKRGSVRAPOVYVLPPEEMTK 244
Qy      347 QVSLCLVKGFPSPDIAYVESNGOPENNYKTPVLDSDGSFFLYSKLTVDSRMQGN 406
Db      245 QVTLCTQMTDFWPEBIDIVYETWNGTELNYKTEPVLDSGSYFMYSKLRVKKMVERN 304
Qy      407 VFSCSVMEALHNNYTKSLSLSPG 431
Db      305 SVSCSVVMEGLHNNHTTSFSPG 329

```

RESULT 36

GMSA

```

Ig gamma-2a chain C region, secreted form (allele b) - mouse
C.Species: Mus musculus (house mouse)
C.Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999
C.Accession: A02153; A32656
R.Schreier, P.H.; Botwell, A.L.M.; Mueller-Hill, B.; Baltimore, D.
Proc. Natl. Acad. Sci. U.S.A. 78, 4495-4499, 1981
A.Title: Multiple differences between the nucleic acid sequences of the IgG2a(a) and IgC
A.Reference number: A02153; MUID:82037861; PMID:6170065
A.Accession: A02153
A.Molecule type: mRNA
A.Residues: 1-335 <SCH>
A.Cross-references: GB:J00479
A.Experimental source: Strain C57BL/6
R.Dognin, M.J.; Laureys, M.; Stroberg, A.D.
Proc. Natl. Acad. Sci. U.S.A. 78, 4031-4035, 1981
A.Title: Multiple amino acid substitutions between murine gamma 2a heavy chain Fc region
A.Reference number: A32656; MUID:82037777; PMID:6794027
A.Accession: A32656
A.Molecule type: protein
A.Residues: 118-267; 'E', 269-328; 'G', 330-334 <DOG>
C.Comment: Iys-335 is removed posttranslationally.
C.Comment: The sequence differs from that of the allele, from BALB/c mice, at 15% of t
C.Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C.Superfamily: immunoglobulin C region; immunoglobulin homology
C.Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
F:20-84/Domain: immunoglobulin homology <IM1>
F:98-118/Region: hinge
F:142-211/Domain: immunoglobulin homology <IM2>
F:248-315/Domain: immunoglobulin homology <IM3>
F:15/Disulfide bonds: interchain (to light chain) #status predicted
F:27-82,149-209,255-313/Disulfide bonds: #status predicted

```


F108,117/Disulfide bonds: Interchain (to heavy chain) #status predicted
F185/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.6%; Score 801; DB 1; Length 335;

Best Local Similarity 50.2%; Pred. No. 1.5e-40;

Matches 166; Conservative 52; Mismatches 77; Indels 36; Gaps 9;

QY 124 GLTANSDT--HLQG---QSLTLTLESPGSS-----PSVOCRSPRGKNIQGGKTLVSQ 173
DB 17 GTTSSVTLGLGVKGFPEPTLTLWNSGSLSGVHTFPAL-----LQSG-LYTLSS 66
QY 174 LELQDSGTW-----TCTV-----LQNKQVEFKIDIV--PCPAPEKSCDKHTCPRL 220
DB 67 SVTVSNMTPSQTITCNVAHPASSTKVCKIEPRVITQNPCE---PQGRVPCAPDL 123
QY 221 GGPVFLFPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFMWYDGVENNAKTKRREQ 280
DB 124 GGPVFLFPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFMWYDGVENNAKTKRREQ 183
QY 281 YNSTRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPOVYTLPPSR 340
DB 184 YNSTRVVSVLPIQHQMDSGKEFKCKVNNALPAPIEKTIISKAKGQPREPOVYTLPPSR 243
QY 341 DELTKNQVSLTCLVKGFIPSDIAVEESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKS 400
DB 244 EEMTKKEFSLTCMTITGFLPAEIAVDMWTSNGRTEQNKATATVLDSDGSFFLYSKLTVDKS 303
QY 401 RMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 431
DB 304 TWERGSLEFACSVNHEVNLHNLTKTISRSLG 334

RESULT 37
S01321

Ig gamma-2b chain precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 23-Jul-1999

C/Accession: S01321

R/de Waale, P.; Feys, V.; van de Voorde, A.; Mollemans, F.; Fiers, W.

Eur. J. Biochem. 176, 287-295, 1988

A/Title: Expression in non-lymphoid cells of mouse recombinant immunoglobulin directed at

A/Reference number: S01320; MUID:88329081; PMID:3138116

A/Accession: S01321

A/Molecule type: mRNA

A/Residues: 1-475 <DB1>

A/Cross-references: EMBL:X13188; NID:651780; PID:CA31580.1; PID:g51781

A/Note: this sequence was determined from the differentiated gene

C/Superfamily: Immunoglobulin C region; Immunoglobulin homology

C/Keywords: Immunoglobulin

F1-19/Domain: signal sequence #status predicted <SIG>

F1-20-475/Product: Ig gamma-2b chain #status predicted <MAT>

F159-223/Domain: immunoglobulin homology <IMM>

Query Match 29.4%; Score 793.5; DB 2; Length 475;

Best Local Similarity 40.0%; Pred. No. 6.5e-40;

Matches 187; Conservative 58; Mismatches 135; Indels 87; Gaps 14;

QY 30 LGKKGDVTELCTAS--QKSGIHFWMKNSNOIKILNGG-----SFTTKPSKLNBRAD 81
DB 30 LARPCASVLSCKASGYTLTISGISW-----VKQRTGGLEWIGETIYFGSGNSYFNEFK 84
QY 82 SRRSLW---DQGNPFLIKNLKIEDSDTYICEVEDQKEVOLLVFGLTANSDTHLQOQS 138
DB 85 GKATLTVDKSSSTAYLHLSLSLSDSNAVYFCA---GPRQVGLLPRG-----YNGQG 132
QY 139 LTLTLESPGSSPSVQ-----CRSPRGKNI-----QGGKTLV--- 171
DB 133 TLVTASAKTTPPSVYPLAPGCGDTTGSSTLGLGVKGFPESTVTVWNSGLSSVATF 192
QY 172 -----SQLELDQSGTW-----TCTVLQ--NOKQVFEKID-----LYPCAPE 206
DB 193 PALIQSGLYTWSSSVTVPSQVTCVVAHPASSTTVDKLEPSPGTSTINPC--- 249

QY 207 PKGCDKHTTC--PELLGPGSVFLFPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFMWY 264
DB 250 --PKECHKCAPRLLEGPSVFIFPPNKKOVLMSLTPKVTCCVVVDVSEDDPDQISWFFV 307
QY 265 DGEVHNAKTKPREEQYNSTIRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKA 324
DB 308 NNVEVLTAQOTQTHREDNSTIRVVSALPIQHODMSKGEFKCKVNNKDLPAPIERTISKI 367
QY 325 KGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFIPSDIAVEESNGQPENNYKTTPVLD 384
DB 368 KGIVAPQVYTLSPPELRSKDVSLTCLVAGSPEDISVEWTSNGHTEENYKDTAPLVD 427
QY 385 SDGSFFLYSKLTVDKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 431
DB 428 SDGSFFLYSKLTNMKTSKMEKTDSPSCVNRHGLKNYLLTKTISRSPG 474

RESULT 38
PS0019

Ig gamma-2a chain C region - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Jul-1999

C/Accession: PS0019; D25941

R/Bruggemann, M.

Gene 74, 473-482, 1988

A/Title: Evolution of the rat immunoglobulin gamma heavy-chain gene family.

A/Reference number: PS0017; MUID:89232738; PMID:3349946

A/Accession: PS0019

A/Molecule type: DNA

A/Residues: 1-322 <BRU>

R/Bruggemann, M.; Free, J.; Diamond, A.; Howard, J.; Cobbold, S.; Waldmann, H.

Proc. Natl. Acad. Sci. U.S.A. 83, 6075-6079, 1986

A/Title: Immunoglobulin heavy chain locus of the rat: striking homology to mouse antibody

A/Reference number: A25941; MUID:86287397; PMID:3016742

A/Accession: D25941

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 216-322 <BR2>

C/Genetics:

A/Intons: 98/1; 109/1; 216/1

C/Superfamily: Immunoglobulin C region; immunoglobulin homology

C/Keywords: Immunoglobulin

F1-20-84/Domain: immunoglobulin homology <IMM>

Query Match 29.3%; Score 792.5; DB 2; Length 322;

Best Local Similarity 51.0%; Pred. No. 4.7e-40;

Matches 160; Conservative 47; Mismatches 72; Indels 35; Gaps 8;

QY 133 LQGG---QSLTLTLESPGSS-----PSVOCRSPRGKNIQGGKTLVSQLELDQSGTW-- 182
DB 28 LVKGVPPEPTVTVWNSGALSSGVHTFPAY-----LQGLYTLSSVTV--PSSTWSS 77
QY 183 ----TCTVLQ--NOKQVEFKIDIVPCPAPEKSCDKHTCPBELLAGPSVFLFPKPKDTLM 237
DB 78 QAVTCVNAHPASSTKVCKIVPREC---NPGCTGSGSV-----SSVTFPPPKTKDVL 127
QY 238 ISRTPEVTCVVVDVSHEDPEVKFMWYDGVENNAKTKPREBOYNSTIRVSVLTVLHQD 297
DB 128 ITLTPKVTCCVVVDISQNDPEVRFSWFTDVEVHTAQHAPEKQNSLTRSSELPIVHRD 187
QY 298 WLNKGEYKCKVSNKALPAPIEKTIISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGF 357
DB 188 WLNKGTFCCKVNSGAPAPAPIEKTSIKREGTRPGQVYTVMAPKEMQSGVSTICWVKG 247
QY 358 YPSDIAVEESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRMOQGNVFCSCVMHEAL 417
DB 248 YPDIYTEMKNGQPCQENNYKTTPMTDSDGSFFLYSKLTVMKKEKTMQCGNTFTCSVLHGL 307
QY 418 HNNHYTOKSLSLSPG 431
DB 308 HNNHTEKSLSHSPG 321

RESULT 39

RMW074

T-cell surface glycoprotein CD4 - rhesus macaque

N.Alternate names: T-cell surface antigen T4/Leu 3

C.Species: Macaca mulatta (rhesus macaque)

C.Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999

C.Accession: C32722

R.Camerini, D.; Seed, B.

Cell 60, 747-754, 1990

A.Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi

A.Reference number: A32722; MUID:90182664; PMID:2107024

A.Accession: C32722

A.Molecule type: mRNA

A.Residues: 1-432 <CAM>

A.Cross-references: GB:M31134

C.Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells

C.Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C.Keywords: duplication; glycoprotein; T-cell; transmembrane protein

F.1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MNT>

F.1-371/Domain: extracellular #status predicted <EXT>

F.9-86/Domain: immunoglobulin homology <IM1>

F.111-161/Domain: immunoglobulin homology #status atypical <IM2>

F.180-293/Domain: immunoglobulin homology <IM3>

F.296-347/Domain: immunoglobulin homology <IM4>

F.372-395/Domain: transmembrane #status predicted <TM>

F.396-432/Domain: intracellular #status predicted <INT>

F.16-84,130-159,303-345/Diulfide bonds: #status predicted

F.271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.2%; Score 790; DB 1; Length 432;

Best Local Similarity 87.4%; Pred. No. 9.4e-40;

Matches 152; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

Qy 27 KYVLGGKGTVEITCTASQKSIQFHWKNSNQIKINGOSPLTKGPSLNDPADRSRL 86

Db 2 KYVLGGKGTVEITCTASQKSIQFHWKNSNQIKINGOSPLTKGPSLNDPADRSRL 61

Qy 87 WPOGNPLLIKNLKIEDSDTYICEVEDKEVQLVFGILTANSDTHLGGSLTLTLSP 146

Db 62 WPOGNPLLIKNLKIEDSDTYICEVEDKEVQLVFGILTANSDTHLGGSLTLTLSP 121

Qy 147 PSSSPVQCRSPKKNIOGKTLVSQLELDSDGTCTVTLNOKKVEFKIDIV 200

Db 122 PSSSPVQCRSPKKNIOGKTLVSQLELDSDGTCTVTLNOKKVEFKIDIV 175

RESULT 40

S06611

Ig gamma-2 chain C region (clone 32.2) - bovine (fragment)

C.Species: Bos primigenius taurus (cattle)

C.Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 16-Jul-1999

C.Accession: S06611; B31303

R.Symons, D.B.A.; Clarkson, C.A.; Beale, D.

Mol. Immunol. 26, 841-850, 1989

A.Title: Structure of bovine immunoglobulin constant region heavy chain gamma 1 and gamma

A.Reference number: S06610; MUID:90087956; PMID:2513487

A.Accession: S06611

A.Molecule type: DNA

A.Residues: 1-327 <SYM>

A.Cross-references: EMBL:X16702

A.Note: the sequence was determined from the germline gene

C.Genetics:

A.Gene: Ig CH gamma 2

A.Introns: 99/1; 112/1; 219/1

C.Superfamily: immunoglobulin C region; immunoglobulin homology

C.Keywords: glycoprotein; immunoglobulin; membrane protein

F.20-85/Domain: immunoglobulin homology <IM>

F.175/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 28.6%; Score 774; DB 2; Length 327;

Best Local Similarity 57.0%; Pred. No. 6e-39;

Matches 146; Conservative 41; Mismatches 55; Indels 14; Gaps 4;

Qy 181 TWTCTVLQ--NOKVEFKIDI-VPCPAPEPKSCDTHTCPELLGSPVFLPPEPKDTLM 237

Db 80 TTTCAVAPASSTKTDKAVGVSSDCSKRNQNCVRE-----PSVFIFFPKKDTLM 130

Qy 238 ISRTPEVTCVVDVSHEDPEVKFNMYVDGVEVNAKTKPREQYNSTRVSVLTVLHOD 297

Db 131 IGTPEVTCVVDVSHEDPEVKFNMYVDGVEVNAKTKPREQYNSTRVSVLTVLHOD 190

Qy 298 WINGEYCKVSNKALPAPIETISKAGQPREPVYTLTPSRDELTKQVSLTCLYKGF 357

Db 191 WINGEYCKVSNKALPAPIETISKAGQPREPVYTLTPSRDELTKQVSLTCLYKGF 250

Qy 358 YPSDIAYEWESNGQ--PENNYKTPPVLDSDGSFFLYSKLTYDKSRMOQGNFSCVME 415

Db 251 YPEDVDVAKQRPQROSEBKTITTPOLDADNSYFLYSKLRVDRNSMQRGDTTCVME 310

Qy 416 ALHNHYTKSLSLSPG 431

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

Qy 311 ALHNHYTKSLSLSPG 326

Db 311 ALHNHYTKSLSLSPG 326

F:322-372/Domain: immunoglobulin homology <IMM>

Query Match 22.3%; Score 601.5; DB 2; Length 459;
Best Local Similarity 41.9%; Pred. No. 1.6e-28;
Matches 153; Conservative 47; Mismatches 98; Indels 67; Gaps 11;

```
OY 1 MNRGVPRHLLVLTQALLPAATQGNKVLGKGGDTVELTCTAQQKSIOFHMKNSNOIK 60
DB 1 MNRRIYFOCLLVLLELPAATWGTAVRGKAGAVELPCOSSQKRSVFNWIKANQVK 60
OY 61 ILNGG-----SFLTGPRSKLNDRADRSRLMDQGFPLIKLKLESDDTTICEVEDQKE 116
DB 61 ILNGGSSSSSSFWLKGNSPLSNRVESKKNMMDQGFPLIKLRRMDSDGTICEGDKKM 120
OY 117 EVOLLVFLGLTANSPDTHLLQGSLLTLLESPPSSPVOCSPRGKNIQGGKTLVSQLEL 176
DB 121 EVELLVFLTLNPNTRLHGOSLTLTLBGPVSGSVQWKSPEKNIITGFTCSMPKRL 180
OY 177 QDSGTWCTV-LQNGKVEFKIDIVPCAPBPKSCDKTHTCPBELLGSGVFLFPKPKDT 235
DB 181 QDSGTWSCHLSFQDNKLELDIKIIVLGFPKASA-----TYKKKEGEVERSPF----- 229
OY 236 LMISTPEVTGVVDVSHEDDEV--KFWWYVDGVVHNAKTKPREEQYNSTRVVSULTV 293
DB 230 -----LNFEDSLSGELMWQVDGAS----- 249
OY 294 LHQDWLNGKYEKCVKS-NKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNOVSLT 351
DB 250 SAQSVSVSLSDRKVSQKILP--DLKIQMSKGLPLS---LTLPLALHRYAGSGLSLT 303
OY 352 CLVKG 356
DB 304 -LDKG 307
```

RESULT 43

S30193
T-cell surface glycoprotein CD4 - dog
C/Species: Canis lupus familiaris (dog)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C/Accession: S30193
R/Milte, K.F.; Conner, G.E.; Mitcz, D.H.; Alejandro, R.
Biochim. Biophys. Acta 1172, 315-318, 1993
A/Title: Primary structure of the canine CD4 antigen.
A/Reference number: S30193; MUID:93192324; PMID:791632
A/Accession: S30193
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-432 <ML>
A/Cross-references: EMBL:X68565; NID:q288652; PIDN:CA837664.1; PID:94467377
C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
C/Keywords: glycoprotein
F:202-311/Domain: immunoglobulin homology <IMM>

Query Match 21.5%; Score 581.5; DB 2; Length 432;
Best Local Similarity 32.8%; Pred. No. 2.3e-27;
Matches 170; Conservative 67; Mismatches 152; Indels 129; Gaps 19;

```
OY 12 LVLTALLPATQGNKVLGKGGDTVELTCTAQQKSIOFHMKNSNOIKILNGGSFLTK 71
DB 1 LMLLVMLPAVTPVNEVVLGKAGDAVELPCQTSQKNIHFWRDSSMWQIILNGGSEFTV 60
OY 72 GPKLNDRADRSRLMDQGFPLIKLKLESDDTTICEVEDQKE 127
DB 61 GSKLKHVESKKNLMDQGFPLIKLESDDTTICEVEDQKE 119
OY 128 -----NSDTLLQGSLLTLLESPPSSPVOCSPRGKNIQGGKTLVSQLELQDSGTW 182
DB 120 GSSGSSNIRLLQGSLLTLLENPSGSSPVQWKGPKMSHGQGNLSLSPBELQDSGTW 179
OY 183 TCTVLQNGKVEFKIDIVPCAPBPKSCDKTHTCPBELLGSGVFLFPKPKDTLMISRTF 242
DB 180 TCIISQSKVEFNIINVLVLAFA--QKVSNTFYARE--GDQVEFSFP----- 221
```

```
OY 243 EVTCVVVDVSHEDDEV--KFWWYVDGVVHNAKTKPREEQYNSTRVVSULTVLDHODVLN 300
DB 222 -----LSEEDNVLGELRMQAGAS-----SSLL-----WIS 248
OY 301 GKEXYKCVSNKALPAPIEKTISKAKGPRE--POVYTLPPSRDELTKNOVSLTCLVKGFY 358
DB 249 FTLENRKLSMKEAALPL-----KLQKESLPLRFTLPQVLSRYAGSGIITLNLAKGTL 301
OY 359 PSDIAVEMESNGQENNYKTPPLVDSGSEFLYSKLTVDKSRQGNVFCSSVHEALH 418
DB 302 YQEV-----NLVWMASSQNNL--TCEVLGP--- 326
OY 419 NHYQKSLSPGQLDETCAEADGELDGLMTDPPRASALPAPFGSALPDPQ---TA 475
DB 327 ---TSPBELTIS--LNLGEQAKVSK--QOKLVWVVDPEGCT-----NQCLLSDKDKVLLA 374
OY 476 SAL--PDPASALPALA-----VISFLGLGLGVAC 506
DB 375 SSLVSSPVVYKSWPKFLATLGLLGLLGLCVFC 412
```

RESULT 44

A36040
Ig heavy chain V-III region (ART) - human (fragments)
C/Species: Homo sapiens (man)
C/Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 21-Jan-2000
C/Accession: A36040
R/Bultz, M.; Weiss, D.T.; Solomon, A.
Proc. Natl. Acad. Sci. U.S.A. 87, 6542-6546, 1990
A/Title: Immunoglobulin heavy-chain-associated amyloidosis.
A/Reference number: A36040; MUID:90370821; PMID:2118650
A/Accession: A36040
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-218 <EUL>
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: heterotetramer; immunoglobulin
F:132-199/Domain: immunoglobulin homology <IMM>

Query Match 21.3%; Score 574.5; DB 2; Length 218;
Best Local Similarity 77.5%; Pred. No. 2.6e-27;
Matches 110; Conservative 3; Mismatches 12; Indels 17; Gaps 1;

```
OY 290 VLTVDHMDWNGKYEKCVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNOVS 349
DB 94 VAAALLTDWNGG-----FLVTVSSQPREPOVYTLPPSRDELTKNOVS 136
OY 350 LTCLVKGYPSPDIADVEMESNGQENNYKTPPLVDSGSEFLYSKLTVDKSRWQGNVFS 409
DB 137 LTCLVKGFPSPDIADVEMESNGQENNYKTPPLVDSGSEFLYSKLTVDKSRWQGNVFS 196
OY 410 CSVWHEALNHNYTQKSLSPG 431
DB 197 CSVWHEGLNHNYTQKSLSPG 218
```

RESULT 45

S69340
Ig heavy chain VHIII-D-JH-CH3 region - human
C/Species: Homo sapiens (man)
C/Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 21-Jan-2000
C/Accession: S69340
R/Khamlich, A.A.; Ancouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995
A/Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A/Reference number: S69339; MUID:95262687; PMID:7744049
A/Accession: S69340
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-249 <KTA>
A/Cross-references: EMBL:X81696
C/Superfamily: immunoglobulin C region; immunoglobulin homology

F1162-229/Domain: immunoglobulin homology <IMM>

Query Match 21.2%; Score 572.5; DB 2; Length 249;

Best Local Similarity 62.9%; Pred. No. 4e-27; Indels 37; Gaps 3;

Matches 117; Conservative 9; Mismatches 23; Indels 37; Gaps 3;

266 GVEVNAKTPREEDQNSTYRVSVLT-----VLHDMWNGKVKCVSN 310
 DB 80 GDSVGRFTISRDNSNOLXLMNSTLADTAVYCAKGVSVIH----- 124
 QY 311 KALPAPIE-----KTSKAKGPREPOVYTLPPSRDELTKNOVSLTLVKGFPYSDIAVE 365
 DB 125 --LPSAIHFHWGQSTRYIVSSGQPREPOVYTLPPSRREEMTKNOVSLTLVKGFPYSDIAVE 182
 QY 366 WESNOGPENNYKTPPVLDSDGSFLYKSLTYDKSKMOGQNVFSCSVMEALHNHYTKS 425
 DB 183 WESNOGPENNYKTPPVLDSDGSFLYKSLTYDKSKMOGQNVFSCSVMEALHNHYTKS 242
 QY 426 LSLSPG 431
 DB 243 LSLSPG 248

RESULT 46

S14236

IG gamma-1 chain C region (15C5) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S14236

R:Vandamme, A.M.; Buleen, F.; Benhar, H.; Nelles, L.; Lijnen, R.H.; Collen, D.

Eur. J. Biochem. 192, 767-775, 1990

A:Title: Construction and characterization of a recombinant murine monoclonal antibody C

A:Reference number: S14236; MUID:91006173; PMID:2209622

A:Accession: S14236

A:Molecule type: mRNA

A:Residues: 1-152 <VAN>

A:Cross-references: EMBL:X56393; NID:951617; PIDN:CAA39804.1; PID:951618

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F165-132/Domain: immunoglobulin homology <IMM>

Query Match 20.9%; Score 566; DB 2; Length 152;

Best Local Similarity 61.6%; Pred. No. 5.4e-27;

Matches 93; Conservative 33; Mismatches 25; Indels 0; Gaps 0;

QY 281 YNSTYRVSVLTVLHDMWNGKVKCVSNKALPAPIEKTISKAKGPREPOVYTLPPSR 340
 DB 1 FNSTPVSVELPIHMDWNGKFKCRVNSAAPPAPIEKTISKAKGPRAPQVYTIIPPK 60
 QY 341 DELTKNOVSLTLVKGFPYSDIAVEWESNOGPENNYKTPPVLDSDGSFLYKSLTYDKS 400
 DB 61 EGMADDKVSLTCMIDTFPEDITVEMQWNGQPAENYKNTQPIMTNGSYFYVSKLVOKS 120
 QY 401 RWQGNVFCSCVMEALHNHYTKSLSLSPG 431
 DB 121 NWEAGNTFCSVLHREGLHNHHTKSLSHSPG 151

RESULT 47

A27449

T-cell surface glycoprotein CD4 precursor - rat

N:Alternate names: W3/25 antigen

C:Species: Rattus norvegicus (Norway rat)

C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Jan-2000

C:Accession: A27449; A35433

R:Clark, S.J.; Jeffries, W.A.; Barclay, A.N.; Gagnon, J.; Williams, A.F.

Proc. Natl. Acad. Sci. U.S.A. 84, 1649-1653, 1987

A:Title: Peptide and nucleotide sequences of rat CD4 (W3/25) antigen: evidence for deriv

A:Reference number: A27449; MUID:87175535; PMID:3104900

A:Accession: A27449

A:Molecule type: mRNA

A:Residues: 1-457 <CLA>

A:Cross-references: GB:M15768; NID:9203387; PIDN:AAA0901.1; PID:9203388

R:Davis, S.J.; Ward, H.A.; Puklavec, M.J.; Willie, A.C.; Williams, A.F.; Barclay, A.N.
 J. Biol. Chem. 265, 10410-10418, 1990
 A:Title: High level expression in Chinese hamster ovary cells of soluble forms of CD4 T
 A:Reference number: A35433; MUID:90285164; PMID:2113054
 A:Contents: annotation
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
 C:Keywords: glycoprotein; membrane protein; surface antigen
 F:219-300/Domain: immunoglobulin homology <IMM>

Query Match 18.3%; Score 495; DB 2; Length 457;

Best Local Similarity 32.8%; Pred. No. 3.3e-22;

Matches 137; Conservative 57; Mismatches 132; Indels 92; Gaps 14;

QY 1 MNRGVPFRHL--LVLQALALPAATQGNKVVLGKGDVETLCTASOKKSIOFHKNSNQ 58
 DB 1 MCRGSPFHLPLPLQLQSLKLVVQGTVLVKGSGSAELCESTSRSSAFWAKSSDQ 60
 QY 59 IKILGNQGSFLTKGSPSKNDRASSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEV 118
 DB 61 KTLIGYKMKLLIKGSLLEYSRDSRKNAMERGSFLLINKLRMEDSOTYVCELENKKEEV 120
 QY 119 QLVFGLTANSDTHLLQGSLTLTLFS--PSSSPSVQCRSPRGKNIQGGKTLVSQLEIQ 177
 DB 121 ELWVFRVTFNPGTRLLQGSLLTLIDSNPKVSDPIECKHKSSNIVKDSKAFSTSLRIQ 180
 QY 178 DSGTTCVTLQNKQKVER--KIDIVPCAPAPKSCDKHTHCELLGFSVLPFPKPKDT 235
 DB 181 DSGINCTVTLNOKHGSFDMKLSVL-----GFASTITAYKSGESAESFP----- 227
 QY 236 IMISRTPEVTCVVDVSHEDPEKFNWYVDGEVNAKTKPREEDQNSTYRVSVLTVLH 295
 DB 228 -----LNLGESSLQGLRW-----KAKKAPS----- 249
 QY 296 QDMWNGKVKCVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNOVSLTLVKG 355
 DB 250 QSMITFSLKQKVS-----VQKSTSNPKFQLSE-----TLQI- 285
 QY 356 GFYPSDIAVEWESNOGPENNYKTPPVLDSDGSFLYKSLTYDKSKMOGQNVFSCSV 413
 DB 286 ----QVSLQFAGSG---NLTLT---LDR-GILVEVNLVWKVTPQDSNTLTCEVM 331

RESULT 48

RMKST4

T-cell surface glycoprotein CD4 precursor - mouse

N:Alternate names: T-cell differentiation antigen L3T4; T-cell surface antigen T4/Leu 3

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A02110; A26038; A39893; A39955; I5464; I69018; A47642

R:Tourville, B.; Gorman, S.D.; Field, E.H.; HunnKapiller, T.; Parne, J.R.

Science 234, 610-614, 1986

A:Title: Isolation and sequence of L3T4 complementary DNA clones: expression in T cells

A:Reference number: A02110; MUID:87018845; PMID:3094146

A:Accession: A02110

A:Molecule type: mRNA

A:Residues: 1-457 <TOU>

A:Cross-references: GB:M13816; NID:9192070; PIDN:AAA37267.1; PID:9309112

R:Litman, D.R.; Gettner, S.N.

Nature 325, 453-455, 1987

A:Title: Unusual insertion in the immunoglobulin domain of the newly isolated murine CD4 (L

A:Reference number: A26038; MUID:87115821; PMID:3027575

A:Accession: A26038

A:Molecule type: mRNA

A:Residues: 1-457 <LIT>

A:Cross-references: GB:X04836; NID:950353; PIDN:CAA28539.1; PID:950354

R:Gorman, S.D.; Tourville, B.; Parne, J.R.

Proc. Natl. Acad. Sci. U.S.A. 84, 7644-7648, 1987

A:Title: Structure of the mouse gene encoding CD4 and an unusual transcript in brain.

A:Reference number: A39893; MUID:88041159; PMID:2823269

A:Accession: A39893

A:Molecule type: DNA

A:Residues: 1-25, 'E' 27-457 <GOR>

A:Cross-references: GB:M17080; GB:003003; NID:9192515; PIDN:AAA37402.1; PID:9387124

R.Maddon, P.J.; Molineaux, S.M.; Maddon, D.E.; Zimmerman, K.A.; Godfrey, M.; Alt, F.W.; Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159, 1987
 A>Title: Structure and expression of the human and mouse T4 genes.
 A:Reference number: A39955; MUID:88097446; PMID:3501122
 A:Accession: A39955
 A>Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 25-457 <MAD>
 A>Note: the cited Genbank accession number, J03564, is not in release 101.0
 R.Parnes, J.R.; Hunkapiller, T.
 Immunol. Rev. 100, 109-127, 1987
 A>Title: L3T4 and the immunoglobulin gene superfamily: New relationships between the imm
 A:Reference number: 154564; MUID:88152875; PMID:3326818
 A:Accession: 154564
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-457 <RES>
 A:Cross-references: GB:M36850; NID:g198670; PIDN:AAA39401.1; PID:g198671
 A:Accession: 159018
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 208-318 <RE2>
 A:Cross-references: GB:M36851; NID:g198672; PIDN:AAA39402.1; PID:g554183
 R.Clackson, B.U.; Tsagaratos, J.; Kitzbaum, L.; Maddox, J.; Mackay, C.R.; Brandon, M.; M
 Immunogenetics 23, 129-132, 1986
 A>Title: The L3T4 antigen in mouse and the sheep equivalent are immunoglobulin-like.
 A:Reference number: A47642; MUID:8616694; PMID:3082751
 A:Accession: A47642
 A:Molecule type: protein
 A:Residues: 27-43 <CLA>
 C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells
 C:Genetics:
 A:Introns: 18/1; 74/1; 128/1; 207/1; 319/1; 386/1; 425/3; 448/2
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
 C:Keywords: alternative initiators; duplication; glycoprotein; T-cell; transmembrane pro
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-457/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>
 F:35-114/Domain: immunoglobulin homology <IM1>
 F:139-190/Domain: immunoglobulin homology #status atypical <IM2>
 F:220-301/Domain: immunoglobulin homology <IM3>
 F:241-457/Product: CD4, brain-specific short form #status predicted <BRA>
 F:321-372/Domain: immunoglobulin homology <IM4>
 F:395-419/Domain: transmembrane #status predicted <TM>
 F:420-457/Domain: intracellular #status predicted <INT>
 F:42-112,159-188,328-370/Diulfide bonds: #status predicted
 F:187,298,323,392/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.6%; Score 475; DB 1; Length 457;
 Best Local Similarity 52.8%; Pred. No. 5, 1e-21;
 Matches 104; Conservative 32; Mismatches 55; Indels 6; Gaps 5;

Qy 1 MNRGVPRFH-LLVLQTLALPAATQGNVVGKGGDTVELTCTASOKKSIQFHKNSQI 59
 Db 1 MCRALSLRRLLLQLQSLAVTQKTLVIGKGBSELCSSOKKTIYFTWKFSQR 60

Qy 60 KILNQG-SFLTQK-PSKLRADRSRLDQGNFPLIKLKIETSDTYICEVEDQK 116
 Db 61 KILQHGKGVILIRGSSPSQF-DRFDSKKGAKESFPLIKLMEBSQTYICLENRKE 119

Qy 117 EVOLLVRLTNSDTHLQGSITLTLES-PPGSSPSVQCSPPGKNIQSKTILSVQLE 175
 Db 120 EVELLVFVTPSPGSLIQGSITLTLDNSKSNVNPLETECHGKGVKSVGSKVLSMNL 179

Qy 176 LODSGTWTCTVLOKOK 192
 Db 180 VQSDSPMNCVTTLDOCK 196

RESULT 49
 B46529
 Ig Y heavy chain (7.8S) - duck
 N:Alternate names: Ig gamma chain (7.8S)
 C:Species: Anas platyrhynchos (domestic duck)

C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
 A:Accession: B46529; S20759
 R:Magor, K.E.; Marr, G.W.; Middleton, D.; Wilson, M.R.; Higgins, D.A.
 J. Immunol. 149, 2627-2633, 1992
 A>Title: Structural relationship between the two IY of the duck, Anas platyrhynchos: mo
 A:Reference number: A46529; MUID:93017865; PMID:1401901
 A:Accession: B46529
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-572 <MAG>
 A:Cross-references: EMBL:X65219; NID:962442; PIDN:CAA6322.1; PID:962443
 A:Experimental source: spleen
 A>Note: sequence extracted from NCBI backbone (NCBI:116127)
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: immunoglobulin
 F:37-120/Domain: immunoglobulin homology <IM>

Query Match 15.2%; Score 411.5; DB 2; Length 572;
 Best Local Similarity 30.0%; Pred. No. 4e-17;
 Matches 103; Conservative 55; Mismatches 136; Indels 49; Gaps 10;

Qy 102 EDSPTVCEV---EDQKEVQLVFGITANSQTHLQGSITLTLESPPGSSPSVQCRSP 158
 Db 265 DDSVELLCVITGSPSPPEVEWLDGAPAH-----LVATMRPQ-----REA 306

Qy 159 RGNKIQGKTLVSQLELDQSGTWTCTVLQ-----NQKVEFKIDIVCPAPAPKSCDK 212
 Db 307 GSKTYMATSGTNVSRBWKAGKAFCTCKVGRPAQTGQGHARF-----CRGSAQCS 360

Qy 213 THTCPELLGSPVFLPPPKKDTLMTSRPEVTCVVVDVSHEDPEVKFNMYDGEVANA 272
 Db 361 I-----QIFVVPSP-GSLYIRQDAKVLVNLPSDASLSISWREKSGALRP 408

Qy 273 KTKPREQVSTRVSVLVTHQDMNGEKYCKNKLPAIEKTISAKQPREPQ 332
 Db 409 DPMVLTEHFGTTFASSSLAISTQDWLAGRFCTVGHEDLPVLGKSIAGKATAPY 468

Qy 333 VYTLPSRDLTKQVSLTCLVKGFPYPSDIAVEW--ESNGQPENNYKTPPVLD--SDGS 388
 Db 469 IFPPRPAEELSLAEVLTCLVKGFPQENHVEQWLRHNVPAFAFTTTPPKSPNDGT 528

Qy 389 PFLYSKITLVDKSRWQGNVFCVMEHAIHNYTQSLSPG 431
 Db 529 FFLYSKITVPAKAWQGVSYACVMEHGLPMRFQRLQKTPG 571

RESULT 50
 S04845
 Ig heavy chain precursor - African clawed frog (Xenopus laevis)
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 26-Aug-1999
 A:Accession: S04845; S05695
 R:Amemiya, C.T.; Haïre, R.N.; Litman, G.W.
 Nucleic Acids Res. 17, 5388, 1989
 A>Title: Nucleotide sequence of a cDNA encoding a third distinct Xenopus immunoglobulin
 A:Reference number: S04845; MUID:89345103; PMID:2503814
 A:Accession: S04845
 A:Molecule type: mRNA
 A:Residues: 1-549 <AME>
 A:Cross-references: EMBL:X15114
 R:Litman, G.W.
 submitted to the EMBL Data Library, April 1989
 A:Reference number: S05695
 A:Accession: S05695
 A:Molecule type: mRNA
 A:Residues: 'LC', 3-308, 'H', 310-549 <LIT>
 A:Cross-references: EMBL:X15114; NID:964799; PID:9763031
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: glycoprotein; heterotrimer; immunoglobulin
 F:26-109/Domain: immunoglobulin homology <IM>
 F:281,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 14.9%; Score 402.5; DB 2; Length 549;

Best Local Similarity 29.1%; Pred. No. 1.3e-16;
Matches 118; Conservative 65; Mismatches 134; Indels 89; Gaps 17;

QY 87 WQGG-----NPLIK-----NKIEDSD-----TYICEYE----- 112
DB 171 WNGSGITSGLNKPPALQSGGLPASSQLTILPSDWKAKKSFECVNEHKPTSTKYOKIE 230
QY 113 --DOKEVO---LVFGLTANSDT-----HLLQGSLTLLTLESPGS 149
DB 221 CDDEPEPIEPTVEILLQGPASSKSVLLCLITGYAPSEIKYMWLNGQ---VTNISPSNS 287
QY 150 SPSSVQCRSPRGKNIQGGKTLVSQLELDQSGTWCTVLONOKKVEFKIDIVCPAPEPKS 209
DB 288 KP---CKENG--TFSSRSKVSVPKEDMNSDSYTCVTH-----PASHXTK 329
QY 210 CDKHTTCPELLGGPSVFLFPPPKDITLMSRTPEVTCVVVDVSHEDPEVKENW--YVDGE 268
DB 330 EASTKKCDTATTPKVDVLPSPKD--LVYTKAKYCVISRMASDT-DLTVQMSRSDGK 387
QY 269 VNAKTKPREQVNSTYRVSVLTVLHODMLNGKEKCVSNKALPAPIETKISAKGP 328
DB 388 ALAPDSAP--EKAYDGTFTYKSLTKISPGDMENKQPCNVVHPDLPSPIEKSIQKSDPG 446
QY 339 REPQVYTLPPSRDELTKQVSLTCLVKGFPYPSDIAVWESNG--QPENNYKTPPYLSD 386
DB 447 TEPITLPPSDDELNDPISLICMLKNRPQDIVFVKQGVTLIEDYVMTTTPVLEBE 506
QY 387 GSFPL-YSLTVDKSRMQGNVFCGSMHEALHNHYTKSLSPGLQDETCAEADGEL 431
DB 507 BEGFTSFSKLTITARSQWNRGATYSCI---AAHNTISORDIKNRG 548

RESULT 51

S14683
Ig mu chain precursor, membrane-bound (clone 201) - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 23-Jul-1999
C/Accession: S14683; S08047
R.Friedlander, R.M.; Nussenzweig, M.C.; Leder, P.
Nucleic Acids Res. 18, 4278, 1990
Article: Complete nucleotide sequence of the membrane form of the human Igm heavy chain.
A/Reference number: S14683; MUID:90332450; PMID:2115996
A/Accession: S14683
A/Molecule type: mRNA
A/Residues: 1-627 <FRI>
A/Cross-References: EMBL:X17115; NID:g33450; PIDD:CAA34971.1; PID:g33451
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: immunoglobulin; membrane protein
F.1-15/Domain: signal sequence #status predicted <SIG>
F.16-627/Product: Ig mu chain #status predicted <MAT>
F.34-117/Domain: immunoglobulin homology <IMM>

Query Match 14.4%; Score 389; DB 2; Length 627;
Best Local Similarity 23.9%; Pred. No. 9.6e-16;
Matches 116; Conservative 77; Mismatches 167; Indels 126; Gaps 16;

QY 36 TWELTCTASQ--KKSQIFWKNKSNQIKILGNQSFITKPSKLNDRADSRSLMOQGNP 93
DB 175 SVAVGLADPLPDSITFSMKYKNNSDISSTRG-----FP 209
QY 94 LIIKKLIKEDSDTYICEVEDQKEVQLVFGLTANSPTLL-----QGSLTIT 142
DB 210 SYLRGKTAATSOVLPSKD-----VMQGTDEHVCCKVQHPNKNKKANPLPYI 258
QY 143 LESPSSSPSVQCR-----SPR-----GKNIQGGKTLVSQ 173
DB 259 AELPPLVSVFVPRDQFPGNPRSKSLICQATGFSRQIQVSWLRGKGVSGVTTDQV 318
QY 174 LELQDSG-----TWICTVLONOKKVEFKIDIVCPAPEPKS 212
DB 319 ABAKESGPTTYKVTSTLTIKESDWLSQSMFTCRV--DHKGLTFQGNASSMCVDDDTAIR 376
QY 213 THTCELLGSPVFLFPPPKDITLMSRTPEVTCVVVDVSHEDPEVKENWYVDGEVHNA 272

DB 377 -----VFALPPS--FASIFLTSTLTCLVDTLTYYD--SVTISWTRONGEAVKT 422
QY 273 KTKPREQVNSTYRVSVLTVLHODMLNGKEKCVSNKALPAPIETKISAKGP--REP 331
DB 423 HTNISSEHPNMTFSAVGSACEDDMNSGERFTCVTHTDLPSPILKOTISRKGVALLHRP 482
QY 332 QVYTLPPSRDELTKQVSLTCLVKGFPYPSDIAVWESNGQP--ENNYKTPPYLD--SD 386
DB 483 DVIILPPARQNLRESATITCLVYGFSPADYFQVMQGRQPLSEKVTASPMPEPQAP 542
QY 387 GSFPLYSKLTVDKSRMQGNVFCGSMHEALHNHYTKSLSPGLQDETCAEADGEL 446
DB 543 GRYPHSLITVSEEBMNGETVTCVVAHEALPNRYTERTVDS-----TEGEVSDERGF 597
QY 447 DGLWTT 452
DB 598 ENLMAT 603

RESULT 52

BMMS
Ig epsilon chain C region (version 1) - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999
C/Accession: A02144
R.Liu, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.
Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982
Article: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.
A/Reference number: A02144; MUID:83117774; PMID:6818553
A/Accession: A02144
A/Molecule type: mRNA
A/Residues: 1-388 <LIU>
A/Cross-References: GB:J00476; NID:g194875; PIDD:AAA38085.1; PID:g387220
C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F.1-44/Domain: immunoglobulin homology (fragment) <IM1>
F.81-149/Domain: immunoglobulin homology <IM2>
F.186-254/Domain: immunoglobulin homology <IM3>
F.290-361/Domain: immunoglobulin homology <IM4>
F.10.51.62.133.205.228.332.382/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 14.4%; Score 388; DB 1; Length 388;
Best Local Similarity 28.3%; Pred. No. 6.2e-16;
Matches 119; Conservative 77; Mismatches 138; Indels 86; Gaps 20;

QY 38 ELTCTASQKKSQIFWKNKSNQIKILGNQSFITKPSKLNDRADSRSLMOQGNPLIK 97
DB 22 ELKVTTSQVTS---WGSASK-----NFTCHVTHPPSFNESHRT-----ILVR 59
QY 98 NKIEDSDTYI---CEVEDQKEVQL--LVFGLTAN--SDTHLLQGSLTLLTLESPGS 149
DB 60 PVNITEPFLLELHSCDNPNAFHSITQLCYFIHNLNVSVSWLMDNREITDTL----- 113
QY 150 SPSSVQCRSPRGKNIQGGKTLVSQLELDQSGTWCTVLONOKKVEFKIDIVCPAPEPKS 209
DB 114 AQVLILKE--EGKLASTCCKMLITEQOMMSSESTFTCKV--TSQGVYLYLHATRCPCDHEBR- 169
QY 210 CDKHTTCPELLGGPSVFLFPPPKDITLMSRTPEVTCVVVDVSHEDPEVKENW--YVDGE 268
DB 170 -----GVITYIIPSPDL-LYONGAPKLTCLVVDSESKNVNVTWNOEKTSV 216
QY 262 ---WYVDGEVHNAKTKPREQVNSTYRVSVLTVLHODMLNGKEKCVSNKALPAPI 317
DB 217 SASQVY---TKHN-----NATSIISILPVAWKMDIEGYGCIYDHPFRKPI 263
QY 318 EKTISKANGQPREQVYTLPPSRDELTKQVSLTCLVKGFPYPSDIAVWESNGQENN-- 375
DB 264 VRSITKTPGORSAPDEVVYFPPPEER--SEDKRTLTCLIONFPREDISVQMLDGKLISSQ 322
QY 376 YKTPPYLSDSGS---FLYSKLTVDKSRMQGNVFCGSMHEALHN--HYQKSLSLSPG 431

Db 323 HSTTP-LKNSNGSFFIFSRLEAVAKTLMTQKQFTQVHIALQKPKLEKTISSIG 381

RESULT 53

S15590
Ig heavy chain - human
C/Species: Homo sapiens (man)
C/Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C/Accession: S15590
R/Name: G.A.M.; Kitchingman, G.R.
Nucleic Acids Res. 19, 2427-2433, 1991
A/Title: mRNA transcripts initiating within the human immunoglobulin mu heavy chain enhancer
A/Reference number: S15590; MUID:91252286; PMID:1904154
A/Accession: S15590
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-474 <NEA>
A/Cross-references: EMBL:X58529
A/Note: the authors translated the codon CAA for residue 265 as Glu
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: heterotetramer; immunoglobulin
F/344-415/Domain: immunoglobulin homology <IMM>

Query Match 14.3%; Score 386.5; DB 2; Length 474;
Best Local Similarity 23.7%; Pred. No. 9.7e-16;
Matches 115; Conservative 78; Mismatches 167; Indels 125; Gaps 16;

QY 36 TVELTCTASQ--KKSIOFHKNKNSQIKILGNQGSFLTKGSKLNDRADSRSLMDQGNFP 93
DB 23 SVAAGCLADPFLPDSITFSWKYKNSDISSTRG-----FP 57
QY 94 LIINKLIKIEDSDTYICEVEDQEEVQLVFLGTANSDDLTL-----QGOSLTLT 142
DB 58 SVLNGKRYAATSOVLTPSKD-----VMGTDHVVCKQHPNGNKKVPLVPI 106
QY 143 LESPPGSSPVQCR-----SPR-----GKNIOGKTLVSQSL 174
DB 107 AELPPKAVFVPPRDFGPNRPSKILCOATGSPRQIOVSWLRGKQVSGVTTDQVQA 166
QY 175 ELQDSC-----TWTCVLONOKKEVFKIDVPCAPPEKSCDKT 213
DB 167 EAKSGPTTYVSTLTIKESDMSQSMFTCRV--DHKGLTFQONASSMCVPPDDTAIR- 223
QY 214 HTCELLGSPVFLPPPKDMLMISRTPEVTCVVDVSHEDPEKKNVYDGEVHNAAK 273
DB 224 -----VFALPPS-FASIFLTSTKTLCLVTLDTLTYD-SVTISMTQNOQAVKTH 270
QY 274 TKPREEQNSTYRVSVLTVLHQMUNGKEYKCKVSNKALPAPIEKTISSAKGQP-REPO 332
DB 271 TNISESHNATFSAVGEASICEEDMNSGERFTCTVHTDLPSPKOTISRPKGVALLHRPD 330
QY 333 VYTLPPSDELT-KNOVSLTCLVKGFPSPDIIVAVESNGQP--ENNYKTTIPVLD--SDG 387
DB 331 VYLLPPAEQNLRESATITCLVTFSPADVFVOMORGOPLSPKQVTSAPMEPPQAPG 390
QY 388 SFPLYSKLTIVKSRMOQGNVSCSVMEALHNHYTKSLSPGLQDETCAEQDQDGLD 447
DB 391 RYFAHSILITVSEEMNGETTYTCVAHAELPNRTERTVKSS-----TEGEVSADDEGFE 445
QY 448 GLWTT 452
DB 446 NLMAT 450

RESULT 54

S37768
Ig mu chain C region - human
C/Species: Homo sapiens (man)
C/Date: 12-Feb-1998 #sequence_revision 12-Feb-1998 #text_change 23-Jul-1999
C/Accession: S37768
R/Hatlintrach, N.; Donadel, G.; Sigounas, G.; Nockins, A.L.
Mol. Immunol. 30, 111-112, 1993

A/Title: Comparison of complete nucleotide sequence of the human IgM heavy chain constant
A/Reference number: S37767; MUID:93109369; PMID:8417370

A/Accession: S37768
A/Molecule type: mRNA
A/Residues: 1-453 <HAR>
A/Cross-references: EMBL:X67301; NID:938407; PIDN:CAA77714.1; PID:938408
A/Experimental source: cell line Ab 63
C/Genetics:
A/Map position: 14q32

C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: glycoprotein; heterotetramer; immunoglobulin
F/21-90/Domain: immunoglobulin homology <IMM1>
F/127-199/Domain: immunoglobulin homology <IMM2>
F/237-305/Domain: immunoglobulin homology <IMM3>
F/344-415/Domain: immunoglobulin homology <IMM4>
F/14/Disulfide bonds: interchain (to light chain) #status predicted
F/28-88,134-197,244-303,351-413/Disulfide bonds: #status predicted
F/46,209,272,279,440/Binding site: carbohydrate (asn) (covalent) #status predicted
F/214,452/Disulfide bonds: interchain (to heavy chain) #status predicted
F/291/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 14.2%; Score 385; DB 2; Length 453;
Best Local Similarity 24.9%; Pred. No. 1.1e-15;
Matches 114; Conservative 73; Mismatches 164; Indels 106; Gaps 15;

QY 36 TVELTCTASQ--KKSIOFHKNKNSQIKILGNQGSFLTKGSKLNDRADSRSLMDQGNFP 93
DB 23 SVAAGCLADPFLPDSITFSWKYKNSDISSTRG-----FP 57
QY 94 LIINKLIKIEDSDTYICEVED---QKEVQLVFLGTANSDDLTLQGOSLTLTLESPPGS 149
DB 58 SVLNGKRYAATSOVLTPSKDVMGTDHVVCKQHPNGNKKVPLVPIAELPPKV 113
QY 150 SPSTVQCR-----SPR-----GKNIOGKTLVSQSLQELQDSC- 180
DB 114 SVFPPRDFGFPNRPKSKILCOATGSPRQIOVSWLRGKQVSGVTTDQVQAESGP 173
QY 181 -----TWTCVLONOKKEVFKIDVPCAPPEKSCDKTTCPEL 220
DB 174 TTYKVTSTLTIKESDMSQSMFTCRV--DHKGLTFQONASSMCVPPDDTAIR----- 223
QY 221 GGSVFLPPPKDMLMISRTPEVTCVVDVSHEDPEKKNVYDGEVHNAAKPREEQ 280
DB 224 -----VFALPPS-FASIFLTSTKTLCLVTLDTLTYD-SVTISMTQNOQAVKTH 277
QY 281 YNSTYRVSVLTVLHQMUNGKEYKCKVSNKALPAPIEKTISSAKGQP-REPOVYTLPPS 339
DB 278 PNAITFSVGEASICEEDMNSGERFTCTVHTDLPSPKOTISRPKGVALLHRPDVYLLPPA 337
QY 340 RDELT-KNOVSLTCLVKGFPSPDIIVAVESNGQP--ENNYKTTIPVLD--SDGSFPLYSK 394
DB 338 REQLNLRRESATITCLVTFSPADVFVOMORGOPLSPKQVTSAPMEPPQAPGRYFAHSI 397
QY 395 LTVDKSRMOQGNVSCSVMEALHNHYTKSLSPG 431
DB 398 LTVSEEMNGETTYTCVAHAELPNRTERTVKSSG 434

RESULT 55

S25705
Ig mu chain - sheep
C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C/Accession: S25705
R/Facit, S.; Nau, F.
Mol. Immunol. 29, 829-836, 1992
A/Title: Isolation and sequence of a cDNA coding for the immunoglobulin mu chain of the
A/Reference number: S25705; MUID:92342148; PMID:1635560
A/Accession: S25705
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-592 <PAT>
A/Cross-references: EMBL:X59994; NID:91269; PIDN:CAA42611.1; PID:912670

C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:483-554/Domain: immunoglobulin homology <IMM>

Query Match 14.2%; Score 381; DB 2; Length 592;
Best Local Similarity 27.0%; Pred. No. 1,8e-15;
Matches 123; Conservative 78; Mismatches 153; Indels 102; Gaps 23;

```

QY 35 DTVELCTASQ--KKSIOFMKNSNQIKILNGSFLTKGPSKLNRAISRSLMGGNF 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 161 NYVALGCLADLPFPNSVSSKKNSTVSEKRWTF---PEVLKD-----GLMSASSQ 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 93 PLIKKLIKEDSDTY-ICEVEDK-----EEVQLV-----VEGLTANSDTILLQ 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 211 VALHSSTFOGTGYLVCEVQHKEDVGKGVPRVEVLSPVSVFVPCNS----- 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 137 QSLITLSEPPGSSPVQCR---SPR-----GKNI-----OGGKT--- 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 264 -----LSNGNSKSLICQATDPSPKQISLSWFRDKIVSDISEGQVETVQSSPTTYR 317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 169 -----LSVQLELDQSGTWTCTVLQONOKVEFKIDIPCAPAPKPKSCDKHTPELLGPS 224
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 318 ANSVLTITREMLISGSAVTCQVEHAKETFOKAS-----SSCDATPPSP--IG-- 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 VFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENM-YVDG--VEVNAKTKPREEQY 281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 364 VETIPSPFAD-IFLTSAKLSCLVNMLASYD--GLNISWHQNGKALETTTY----PERIL 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 282 NSTYRVASVLTVLHODMLNGEKYCKVSKNALPAPIEKTISKAGK-QPREPOVYTLPPSR 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 418 NQTFSGRGRASVCSSEMGSEETCTVAHLDLPFPEKSAISPKQVAMRPSPYVLPPIR 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 DELT-KNQVSLTCLVGFYPSDIAVESNGQP--ENNYKTPPVLD--SDGSFLYKSL 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 478 EQSLRESASVTLGVGFAPADVFQVLQKGEVAAASKYVTSAPADPPSAFVHSL 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 396 TVDKSRMQGNVFSQVMHEALNHHTOKSLSPG 431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 538 TVTEEDMSKERTYTCVGHGHALPHMYTERTVDKSTG 573
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 56

MHMM

IG mu chain C region, membrane-bound splice form - human

C:Species: Homo sapiens (man)

C>Date: 29-Jan-1993 #sequence, revision 23-Aug-1997 #text_change 22-Jun-1999

C:Accession: S16510; S09357; S16656; B26243; A02167

R:Dorel, H.; Gillies, S.D.

Nucleic Acids Res. 17, 6412, 1989

A>Title: The complete nucleotide sequence of a human immunoglobulin genomic C-mu gene.

A:Reference number: S09357; MUID:89366650; PMID:2505237

A:Accession: S16510

A:Molecule type: DNA

A:Residues: 433-473 <DOR1>

A:Cross-references: EMBL:X14939

A>Note: the authors translated the codon AAC for residue 445 as Met

A>Note: the sequence of residues 1-432 was assumed to be identical with the correspondin

A:Accession: S09357

A:Molecule type: DNA

A:Residues: 1-432, 'GKPTLVNLSVMSDPTACTY' <DOR2>

A:Cross-references: EMBL:X14940

A>Note: the authors translated the codon AAT for residue 16 as Met

A>Note: secreted splice form

R:Dorel, H.

submitted to the EMBL Data Library, April 1989

A:Reference number: S16656

A:Accession: S16656

A:Molecule type: DNA

A:Residues: 1-39, 'L', '41-432, 'GKPTLVNLSVMSDPTACTY' <DOR3>

A:Cross-references: EMBL:X14940

R:Rabbits, T.H.; Forester, A.; Milstein, C.P.

Nucleic Acids Res. 9, 4509-4524, 1981

A>Title: Human immunoglobulin heavy chain genes: evolutionary comparisons of C-mu, C-del

A:Reference number: A26243; MUID:82059479; PMID:6795593

A:Accession: B26243

A:Molecule type: DNA

A:Residues: 433-436, 'N', '438, 'E', '440-447, 'T', '449-473 <RAB>

A:Cross-references: GB:K01310; NID:9184715; PIDN:AA859422.1; PID:9184734

C:Comment: During differentiation, B lymphocytes switch from expression of membrane-bou

B.

C:Genetics:

A:Gene: GDB:IGHM

A:Cross-references: GDB:120086; OMIM:147020

A:Map position: 14q32.33-14q32.33

A:Introns: 1/1; 105/1; 217/1; 323/1; 433/1; 471/3

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (key

hain disulfide bonds. The IGM subunits associate into disulfide linked pentamers.

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: alternative splicing; glycoprotein; heterotetramer; immunoglobulin; membran

F:21-90/Domain: immunoglobulin homology <IMM1>

F:127-199/Domain: immunoglobulin homology <IMM2>

F:237-305/Domain: immunoglobulin homology <IMM3>

F:344-415/Domain: immunoglobulin homology <IMM4>

F:433-473/Domain: carboxyl-terminal #status predicted <CTS>

F:14/Disulfide bonds: interchain (to light chain) #status experimental

F:18-88,134-197,244-303,351-413/Disulfide bonds: #status experimental

F:46-209,272,279/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:214/Disulfide bonds: interchain (to heavy chain) #status experimental

F:291/Disulfide bonds: interchain (to mu chain in another subunit) #status experimental

Query Match 14.1%; Score 381; DB 1; Length 473;

Best Local Similarity 23.9%; Pred. No. 2.1e-15;

Matches 116; Conservative 77; Mismatches 166; Indels 126; Gaps 17;

```

QY 36 TVELCTASQ--KKSIOFMKNSNQIKILNGSFLTKGPSKLNRAISRSLMGGNF 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 23 SAVAGLADLPDPDSITFSWKYKNSDISSTRG-----PP 57
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 94 LIKKLIKEDSDTYICEVEDQEEVQLVFGLTANSDTILL-----QGSLTYLT 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 58 SVLRGKKAATQVLLPSKD-----VMQSTDEHVCKQVQHPNKNKKNPLPIY 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 143 LESPPGSSPVQCR-----SPR-----GKNIQGGKTLVSQQL 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 107 AELPPKVSFVFPFGFPGNPRKSKLICQATGFSRQIQVSWLRGKQVGSVTTDQVQA 166
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 ELQDSG-----TWTCVLQONOKVEFKIDIVPEPAPEPSCDT 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 167 EKKEGPTTYKYTSLTIKESDWLQGSMTFCV--DHGGLTFQQAASSMCPDODTAIR- 223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 214 HTCPRLGQPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMVYDGVVHNAAK 273
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 224 -----VFALPPS-FASIFLTSTKLTLCLVTLDTLTYD-SVTISWTRNGEAVKTH 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 274 TKPREQYNSTYRVASVLTVLHODMLNGEKYCKVSKNALPAPIEKTISKAGQP-REPO 332
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 271 TNISSEHNAPFSAVGEASICEPDWNSGSRFTCTVTHDLPFLKOTISRPGVALHHRD 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 333 VYTLPPSDDELTKNQVSLTCLVKGFYPSDIAVESNGQP--ENNYKTPPVLD--SDG 387
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 331 VYLLPAPAEQLNLRASATITCLVTFSPADVFQVMQNGQPLSPKRYTSAIMPPOAFG 390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 388 SFFLYSKLTVDKSRMQGNVFSQVMHEALNHHTOKSLSPGLQLETCLEADQDELD 447
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 391 RYFAISILTVSEENMTGETTYC-VAHALLPRKVTERTVDKS-----TBGEVSADBEGR 444
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 448 GLWTT 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 445 NLWAT 449
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 57

S38864

IG epsilon chain C region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-May-2001

F:46,209,272,279,439/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:214,451/Disulfide bonds: interchain (to heavy chain) #status experimental
 F:291/Disulfide bonds: interchain (to mu chain in another subunit) #status experimental

Query Match 13.9%; Score 376.5; DB 1; Length 452;
 Best Local Similarity 24.1%; Pred. No. 3.6e-15;

Matches 112; Conservative 75; Mismatches 156; Indels 121; Gaps 16;

```

QY 36 TVELTQTAG--KKSIOFHMKNSNQIKILGNOSFLLTKPSKLNDAADSRSLMDGNRP 93
DB 23 SYAVGLADFLPDSTTFSMKTKYKNSDISSTRG-----PP 57
QY 94 LIIKLIKIEDSDTYICEVEDOKEVOLLVFGLTANSDFHL-----QGSLTPT 142
DB 58 SVLRGGKVAATSGVLLPSKD-----WVGQDEHVCVKVQHPNKGKNVLPVI 106
QY 143 LSSPPGSPSVOCR-----SPR-----GKNIQGGKTLVSQ 174
DB 107 AELPPKVSFVPRDGFPGNPKSKLICATGFSFPIQVSWLRGKQVSGVTTDQVQA 166
QY 175 ELQDSG-----TWCTVLOQKQKVEFKIDIVPCPAPRPSCDT 213
DB 167 EKESGPTTKYKTSITLTIKESDMLGQSMFTCRV--DHRGLTFQNASMCVDPQDTAIR- 223
QY 214 HTCPBLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAK 273
DB 224 -----VPAIFPS-FASIFLTKSKTKLTCLVTDLTYYD-SVYISWTRONGEAVKTH 270
QY 274 TYPREQVSTYRVSVLTVLVHODWLNKYEKKVSKNKAIPAEIKTISKAKQP-REPQ 332
DB 271 TWISESHPATFSAVGEASICEEDMNSGRFCTVHTDLPSPKQTIIRPKGVALHRD 330
QY 333 VYTLPSRDELTKQVSLTCLVKGFPYSDIAVEMSNQOP--ENNYKTPPLVD--SG 387
DB 331 VYLLPPARQQLMLRESATITCLVTGFSPADVFQVMQROQPSPEKTYVSAPMPEQAIQ 390
QY 388 SFPLYSKLTIVDKSRMOQGVNFSQVMEALHNHYTKSLSLSPG 431
DB 391 RFPANSHLVSEEMWTGETTTC-VNHEALPNRYTERTYDKSG 433

```

RESULT 59

S03186

Ig heavy chain C region - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #ext_change 20-Sep-1999

C:Accession: S03186

R:Haite, R.N.; Shamloct, M.J.; Amemiya, C.T.; Litman, G.W.

Nucleic Acids Res. 17, 1776, 1989

A:Title: A second xenopus immunoglobulin heavy chain constant region isotype gene.

A:Reference number: S03186

A:Accession: S03186

A:Molecule type: mRNA

A:Residues: 1-448 <HA1>

A:Cross-references: EMBL:X13779; NID:G64827; PID:G64828

A:Note: The authors translated the codon TTT for residue 9 as Ser and ATG for residue 34

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 13.8%; Score 374; DB 2; Length 448;
 Best Local Similarity 26.8%; Pred. No. 5e-15;

Matches 117; Conservative 57; Mismatches 144; Indels 118; Gaps 16;

```

QY 80 ADSRSLMDQNFPLIINKLK--IEDSD-----TYICEV---EDQ 114
DB 35 ADPLDIQMDGSIITTKIMRPLVSDVGLYTLSQLTLASBWKSTYCKKVHNYTNT 94
QY 115 KEKVOLLVFGLTANSTHLLQGSLLTLTESPPGSSP----- 151
DB 95 KQEKSLKVLPCMA-----PHYQLFQSPCMQSAISRAGHENINATLDLCTINNEY 145
QY 152 -----SVQCRSPRGKNIOGKTLVSQLELQDSGTW-----TCTVLQN 189

```

DB 146 HQGIKVKWLVNKGQDVSAEASVPTESKTEDGYSSSSQLRIL-KGMNKGQVSCIVTHT 204

QY 190 QKKVERKIDIVCPAPEPSCOKTHCPBLDGGPSFLPPKPKDTLMISRTPEVTCV- 248

DB 205 SSNTTTIANISQC-----TEQCHDNL--QVPLTFPHD-LYFSRNAKVTCLVS 250

QY 249 -----VDVSHEDPEV-KFNMYVDGVEVHNAKTRERQVSTYRVSVLTVLHODMKN 300

DB 251 SKKTIENPISWREKAGNLEFVTEDEPVLIHD-----NGTYSVASILSVCAEDWES 300

QY 301 GKRYCKVSKNKAIPAEIKTISK-AQPREQVYTLTPSRDELTKNQ-VSLTCLVKGFP 358

DB 301 GPKFSCTVASQDLPSPVKTIKFNQEGTPKADVYLLPESAGELIQEWVTLTCTFPTGHN 360

QY 359 PSDIAVEMSNQ---QPENNYKTPPLVLDSDGSFLYSGLTVDKSRMOQGVNFSQVME 415

DB 361 PKEIFQMQGQVSVISEDFPINTVPKPSGEQTYFLYSKLAIPAKMQGDVFTCVVGH 420

QY 416 ALHNHYTKSLSLSPG 431

DB 421 ALPLVITQSDIKSSG 436

RESULT 60

MHMSM

Ig mu chain C region, membrane-bound form - mouse

C:Species: Mus musculus (house mouse)

C:Date: 31-Oct-1980 #sequence_revision 30-Jun-1991 #ext_change 22-Jun-1999

C:Accession: A02167; A37517; B02166

R:Barly, P.; Rogers, J.; Davis, M.; Calame, K.; Bond, M.; Wall, R.; Hood, L.

Cell 20, 313-319, 1980

A:Title: Two mRNAs can be produced from a single immunoglobulin mu gene by alternative

A:Reference number: A02167; MUID:80222874; PMID:6771020

A:Accession: A02167

A:Molecule type: DNA

A:Residues: 436-476 <EAR>

A:Cross-references: GB:J00816; GB:J00444; NID:G52343; PID:CAA24197.1; PID:G52344

R:Rogers, J.; Barly, P.; Carter, C.; Calame, K.; Bond, M.; Hood, L.; Wall, R.

Cell 20, 303-312, 1980

A:Title: Two mRNAs with different 3' ends encode membrane-bound and secreted forms of

A:Reference number: A37517; MUID:80222873; PMID:6771019

A:Accession: A37517

A:Molecule type: mRNA

A:Residues: 410-476 <ROG>

A:Cross-references: GB:V00821; NID:G52355; PID:CAA24202.1; PID:G817972

R:Kawaham, T.; Takahashi, N.; Honjo, T.

Nucleic Acids Res. 8, 3933-3945, 1980

A:Title: Complete nucleotide sequence of mouse immunoglobulin mu gene and comparison with

A:Reference number: A02166; MUID:81076590; PMID:6255422

A:Accession: B02166

A:Molecule type: DNA

A:Residues: 1-435; 'GKPLVNVSLMSPTGTCV' <KAW>

A:Comment: The sequence of residues 1-409 was assumed to be identical with the correspond

C:Comment: During differentiation, B lymphocytes switch from expression of membrane-bound

B:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (lambda)

chain disulfide bonds. In some cases, such as Iga and IgM, the subunits associate into 1-

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobulin

F:21-91/Domain: immunoglobulin homology <IMM1>

F:129-201/Domain: immunoglobulin homology <IMM2>

F:239-307/Domain: immunoglobulin homology <IMM3>

F:346-417/Domain: immunoglobulin homology <IMM4>

F:436-476/Domain: carboxyl-terminal <CTS>

F:14/Disulfide bonds: interchain (to light chain) #status predicted

F:28-89,136-199,246-305,353-415/Disulfide bonds: #status predicted

F:16,211,243,258,281/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:16/Disulfide bonds: interchain (to heavy chain) #status predicted

F:293/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 13.8%; Score 373.5; DB 1; Length 476;
 Best Local Similarity 25.2%; Pred. No. 5.8e-15;

Matches 121; Conservative 87; Mismatches 157; Indels 115; Gaps 23;

QY VELTCTASQ--KKSIOFPMKNSNQIKIINGQSFLLTKGPSKLNDRADSRSLMDQGN-- 92
 DB VAMGLADFLPSTISFTWNYQNTVEVIGQIRFTPT-----LRTGGKTLA 68
 QY 93 ---PLIINKLIKEDSDTY-ICEVEDQKEVOLVLEGLTANSDTHL----LOGQSILTLTLE 144
 DB 69 TSQVLSFKSLIEGDEYLVCKIH-----YG-GKNRDLHPPIPAVAMENNVAVF 117
 QY 145 SPP-----GSSP---SVQCR---SPR-----GKNIOGG----- 166
 DB 118 VPPRDFGSGPARPSKSLICEATNFPKPIITVSMWKDGLVBSGFTTDPVTIENKSGTPQT 177
 QY 167 ----KTLVSOLELQDSGTMCTVQONQKVEFKIDIPCPAPEKSCDKHTCPBLGG 222
 DB 178 YKVISTLTITSELDWNLNVYTCRV--DHKGLTFLKNVSTCAASST-----DIL-- 225
 QY 223 PSVFLPFPKPDITLMISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVNAKTKPREEQYN 282
 DB 226 --TFTIPSPAD-IFLSKANLTLGVSNLATYE-TLINSMSQSGEPLETKIKIMESHFN 281
 QY 283 STYRVSVLTLYHODWLNKGEYKCKVSKALPAPLEKTIISAKGQPRE----POVYTLR 337
 DB 282 GTFSAKGVASVCVEDMNRKEVCTVTRDLPSPCKRISK---PNEVHKHPAVYLLP 337
 QY 338 PSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQ--PENNYKTPPVLD--SDGSFPLY 392
 DB 338 PAREDLNRESATVTCVKGPSPAISIVQMLORGLLQEKVYSAPMEPGAGCFYTH 397
 QY 393 SKLTVDKSRMOQGNVFSQSVNHEALHNHYTQKSLSLSPGLDLETCAEADGELDGLMTT 452
 DB 398 SILFTVEEMNSGETYTCVVGHEALPHLVTERTVDKS-----TDEGVAAEEGFENLWTT 452

RESULT 61

EHRT
 Ig epsilon chain C region - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 17-Dec-1982 #sequence revision 17-Dec-1982 #text_change 16-Jul-1999
 C/Accession: A03442; A03937; A02143
 R/Hellman, U.; Petersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.
 Nucleic Acids Res. 10, 6041-6049, 1982
 A/Title: Structure and evolution of the heavy chain from rat immunoglobulin E.
 A/Reference number: A03442; MUID:83064537; PMID:6292865
 A/Accession: A03442
 A/Molecule type: mRNA
 A/Residues: 1-429 <HEU>
 A/Experimental source: strain LOU/c/Wel, immunocytooma IR2
 R/KindsVogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.
 DNA 1, 335-343, 1982
 A/Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction,
 A/Reference number: A09337; MUID:83182019; PMID:6820340
 A/Contents: myeloma IR162
 A/Accession: A09337
 A/Molecule type: mRNA
 A/Residues: 'N', 169-307, 'L', 309-342 <KIN>
 C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a C/Superfamily: immunoglobulin C region; immunoglobulin homology
 C/Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F/19-80/Domain: immunoglobulin homology <IM1>
 F/118-186/Domain: immunoglobulin homology <IM2>
 F/223-291/Domain: immunoglobulin homology <IM3>
 F/327-398/Domain: immunoglobulin homology <IM4>
 F/46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.7%; Score 370.5; DB 1; Length 429;
 Best Local Similarity 31.5%; Pred. No. 7,7e-15;
 Matches 88; Conservative 62; Mismatches 102; Indels 27; Gaps 10;

QY 160 GKNIOGGKTLISVQLELQDSGTMCTVQONQKVEFKIDIPCPAPEKSCDKHTCPBL 219
 DB 160 GKNIOGGKTLISVQLELQDSGTMCTVQONQKVEFKIDIPCPAPEKSCDKHTCPBL 219

DB 160 GKLAITYSRNLNITQOOWMSESTFCVTSQGE--NYMAHTRRCSDDEPR----- 206
 QY 220 LGGSVFLPFPKPDITLMISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVNAKTKPRE 279
 DB 207 --GVITTIIPSPD-LYENGTPKTLVLDLSEBE-NITVWRERKKSIGSASQSTK 262
 QY 280 QYNSTYRVSVLTLYHODWLNKGEYKCKVSKALPAPLEKTIISAKGQPREPOVYT-LPP 338
 DB 263 HHNNTTSITSLTPVADADWIEGEGYQCRVDPHPKPIVRSITAPGKRSAPVYVFLP 322
 QY 339 SRDELTKNQVSLTCLVKGFYPSDIAVEM--ESNQEPENNYKTPPVLDSDG--FLYS 393
 DB 323 EEEB--VDKRTLTCLIONFFPEDISVQMLQDSKILPKSQHSTTP-LKYNSNORFFIFS 379
 QY 394 KLTVDKSRMOQGNVFSQSVNHEALHN-HYQKSLSLSPG 431
 DB 380 RLEVTKALWITQKQFCTGVHLEALREPRKLERTISKSLG 418

RESULT 62

MHRBM
 Ig mu chain C region, membrane-bound form - rabbit
 C/Species: Oryctolagus cuniculus (domestic rabbit)
 C/Date: 04-Dec-1986 #sequence revision 30-Jun-1991 #text_change 23-Aug-1997
 C/Accession: A02165; A02164
 J/Bernstein, K.E.; Alexander, C.B.; Reddy, E.P.; Mage, R.G.
 J. Immunol. 132, 490-495, 1984
 A/Title: Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain of V-Ha2 a
 A/Reference number: A02164; MUID:84088930; PMID:6418803
 A/Contents: a2 allotype
 A/Accession: A02165
 A/Molecule type: mRNA
 A/Residues: 439-479 <BE2>
 A/Accession: A02164
 A/Molecule type: mRNA
 A/Residues: 1-438, 'GKPTLYNSLIMSDASTCY' <BE2>
 A/Note: the sequence of residues 1-438 was assumed to be identical with the correspondin
 C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a C/Superfamily: immunoglobulin C region; immunoglobulin homology
 C/Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobulin
 F/21-92/Domain: immunoglobulin homology <IM1>
 F/130-202/Domain: immunoglobulin homology <IM2>
 F/242-310/Domain: immunoglobulin homology <IM3>
 F/349-420/Domain: immunoglobulin homology <IM4>
 F/439-479/Domain: immunoglobulin homology <CTS>
 F/14/Disulfide bonds: interchain (to light chain) #status predicted
 F/28-90,137-200,249-308,356-418/Disulfide bonds: #status predicted
 F/46,114,212,261,277,284/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F/219/Disulfide bonds: interchain (to heavy chain) #status predicted
 F/296/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 13.7%; Score 369.5; DB 1; Length 479;
 Best Local Similarity 23.8%; Pred. No. 1e-14;
 Matches 123; Conservative 79; Mismatches 153; Indels 161; Gaps 23;

QY 21 AATGQNVYVCGKGDVTELTCTASQ--KKSIOFPM--KNSNQIKIINGQSFLLTKGPSKL 76
 DB 17 ALTDGNLVANG-----CLARDFLPSSVTFSMSKNSSEI----- 50
 QY 77 NDRADSRSLMDQGNFPLIINKLK-----IBSDTY-ICEVEDQKEVOL 121
 DB 51 -----SSRTV---RTFVVVKRGDKYMATSGVLVPSKVDLQGTETEVLCVKQHSNNRDLR 102
 QY 122 VFGITANSDTHLLOGQLTLTLESPPSSPSVQCR-----SPR-- 159
 DB 103 V-----SFPVDSLELPNVSVFIIPRDSFGSGTRKSRILCOATGSPKQI 147
 QY 160 -----GKNIOGG-----KTLISVQLELQDSGTM-----TCTV-----L 187
 DB 148 SVSMLRGQKVESCVLTLPVDAETKAGAPATFSSMLTTTESQWLSQSLYTCKDHRGCI 207
 QY 188 QNOKKVEFKIDIVCPAPEKSCDKHTCPBLLGSPVFLPFPKPDITLMISRTPEVTCV 247

```

Db      208 FEDKSNVMSSECTSP-----GIQVPIAPSPADT-FLSKSARLICTL 250
Qy      248 VVDVSHDEDEVKFNWVDGVEVHNAKT-----KPREQYNSTRYRVSVLTVLHODMLNCK 302
Db      251 VYDLTLYTG-SLNI-SW-----ASHNGALDTHMNTESHNPATFSAMGASVCAEDWESGE 304
Qy      303 EYKCKVSNKALPAPIEKTISKAKGPPE-POVYTLPPSRDELTKNOVSLTCLVKGFPYS 360
Db      305 QCTCTVTHADLPFLKHTISKSEVAKHPPAYVLPAPAEQLVRESATVTCLVKGFPSPA 364
Qy      361 DIAVESNGQP-ENNYKTPPVLD-SDGSFELYSKLTVDSRWQGNVSCSVMEHA 416
Db      365 DVFVGQORQOPLSBDKYVTSAPAPAPQAPGLYFTHSTLTVEEDMNSGETFTCVVGHBA 424
Qy      417 LNNHTQKSLSPGLQDDETCAEADGELDGLMTT 452
Db      425 LPHMYTERTVDS-----TEGEVGAEEEGFENLMTT 455

```

RESULT 63

```

Ig mu chain C region - dog (tentative sequence)
C:Species: Canis lupus familiaris (dog)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 31-Mar-2000
C:Accession: A93131; A94246; A02169
R:McCumber, L.J.; Capra, J.D.
Mol. Immunol. 16, 565-570, 1979
A>Title: The complete amino-acid sequence of a canine mu chain.
A:Reference number: A93131; MUID:80077682; PMID:117299
A:Contents: myeloma protein Moo
A:Accession: A93131
A:Molecule type: protein
A:Residues: 1-177 <MCC>
R:Wasserman, R.L.; Capra, J.D.
Science 200, 1159-1161, 1978
A>Title: Amino acid sequence of the Cc region of a canine immunoglobulin M: interspecies
A:Reference number: A94246; MUID:78180587; PMID:653360
A:Contents: Moo
A:Accession: A94246
A:Molecule type: protein
A:Residues: 178-450 <WAS>
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) heavy disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 16 C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:120-89/Domain: immunoglobulin homology <IMM1>
F:126-196/Domain: immunoglobulin homology <IMM2>
F:234-302/Domain: immunoglobulin homology <IMM3>
F:341-413/Domain: immunoglobulin homology <IMM4>
F:430-450/Domain: immunoglobulin homology <CTS>
F:114/D:Disulfide bonds: interchain (to light chain) #status predicted
F:127-87,133-194,243-300,348-410/Disulfide bonds: #status predicted
F:50,206,269,276,437/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:211,449/Disulfide bonds: interchain (to heavy chain) #status predicted
F:288/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

```

```

Qy      36 TYELTCTAG--KKSQIHFHKNXNQIKILNGSFLTKPSKLNADRSRLMDQGNP 93
Db      22 TVAMGGLADFLPGSLTFSMKYEBLSAINSTRG-----FP 56
Qy      94 LIIKMLK-----IEDSPTI-CEVE-----DOKEVOLVFGLTANSPTHL 134
Db      57 SVLRGKYATQVFLPSVDIIQGTDEHIVCKYRHSBGRKQKQKVPVLA----- 105
Qy      135 QGOSLTLLSPGSSPSVOCR-----SPR-----GKNIQGGKT 168
Db      106 -----LTL--PPEVSGFIPPRDAFGBPRKSQILCOAGSFGPRQVWSLRDQKQISSGV 157

```

```

Qy      169 LSVQLEIODSG-----TWTC-----TVLQNKKEFKIDIV 200
Db      158 TNEVLAZAKZSGPTTYKVTSMULTIOEDAMLSGVTFCKVEHRLGLFQGNASSM----- 210
Qy      201 PCPAPKPSCDKTHTCPELLGGPSVFLPPPKKDTLMTSRPEVTCVVDVSHEDPEYKE 260
Db      211 -CTSDQPV-----GISIFTLPPS-FASIFMTKSAKLSCLVLDLATVY-SVTI 254
Qy      261 MNVYDGEVHNAKTKPREQYNSTRYRVSVLTVLHODMLNGEYKCKVSNKALPAPIEKT 320
Db      255 SWTRENGALKTHNIISSHPNGTFSAMGEATVCEEWESGQFTCTVTHDLPVLYKOT 314
Qy      321 ISKARG-OPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVESNGQ--PENNY 376
Db      315 ISRPKGVAVHNPSSVVLPPSRQQLDRESATLSCLVTCGSPPDVFWQVQKQPVPPDSY 374
Qy      377 KTPPVLD-SDGSFELYSKLTVDSRWQGNVSCSVMEHALNNHTQKSLSPG--- 431
Db      375 VTSAPMPPEQAPGLYFAHSILTVSEEWNAGETVCAVAHESLPNRVTERSVDSKGTGKPT 434
Qy      432 -----LQDETCAE 440
Db      435 LVNVSILVSDTAGZ 448

```

RESULT 64

```

MHMUR
Ig mu heavy chain disease protein (Bot) - human
C:Species: Homo sapiens (man)
C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 16-Jul-1999
C:Accession: A02163
R:Barnikol-Matanabe, S.; Mihaesco, E.; Mihaesco, C.; Barnikol, H.U.; Hillechmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 365, 105-118, 1984
A>Title: The primary structure of mu-chain-disease protein BOT. Peculiar amino-acid seq
A:Reference number: A02163; MUID:84184186; PMID:6425189
A:Accession: A02163
A:Molecule type: protein
A:Residues: 1-391 <BAR>
C:Comment: This protein has no V region homology or CH1 region.
C:Genetics:
A:Gene: GDB:IGHM
A:Cross-references: GDB:120086; OMIM:147020
A:Map position: 14q32.33-14q32.33
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin; transmembrane pr
F:1-41/Domain: pre-C <VAR>
F:43-391/Domain: Ig mu chain C region, secreted form <IGM>
F:65-137/Domain: immunoglobulin homology <IMM1>
F:175-243/Domain: immunoglobulin homology <IMM2>
F:282-353/Domain: immunoglobulin homology <IMM3>
F:147,210,217,378/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

Qy      110 EYED--QKEVQVLVFGLTANSPTHLIQQOSLTLLSPGSSPSVOCR----- 156
Db      18 EAEDRIIIEEARL-----SGRD--MQVNSQPIAELPPEVSVFVPPRDFGPNRKS 68
Qy      157 -----SPR-----GKNIQGGKTLSVSLLEIODSG-----TWTCVLYON--- 189
Db      69 KLIQATGFSPPQIVSWLRBSKQVSGVTTDEVAEAKESGPTTYKVTSTLIKESDML 128
Qy      190 -QKVEFKIDIVPCPAPKPSCDKTHTCPELLGGPS-----VFLPPPKKDTLMTSRP 242
Db      129 QGSMTCVDRHGLFLQGNASS-----MCGSDQDTAIRVFAIPPS-FASIFLTXT 178
Qy      243 EYTCVAVVDVSHEDPEVKNWVYDVEVHNAKTKPREQYNSTRYRVSVLTVLHODMLNCK 302
Db      179 KTLCLVTLTLTYD-SVTISMTQRQDEAVKTHNIISSHPNATFSAMGASVCAEDWDSDGE 237
Qy      303 EYKCKVSNKALPAPIEKTISKAKGP-REPOVYTLPPSRDELTKNOVSLTCLVKGFPYS 360

```

```

Db      238 RFTCTVHTDLPSPKQITSRPKGVALHRPDYLLPAREQLNRESATITCLVTGFSFA 297
Qy      361 DIAVWESNGOP--ENNKTTPVVD--SDGSFPLXSLTVDKSMQGNFSCSVNHEA 416
Db      298 DVFVWMRGQGLSEPKYVTSAPMPEPOAPKGFYHSLITVSEEMNGETTYCVVAHEA 357
Qy      417 LHNHYTKSLSPG 431
Db      358 LPNRYTERTVDKSTG 372

```

RESULT 65

EMMS

Ig epsilon chain C region (version 2) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Aug-1996

C:Accession: A02145

R:Identi, N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.

EMBO J. 1, 1117-1123, 1982

A:Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison with

A:Reference number: A0966; MUID:84236092; PMID:6329728

A:Accession: A02145

A:Molecule type: DNA

A:Residues: 1-423 <ISH>

A>Note: the sequence was determined from the germline gene

C:Genetics:

A:introns: 91/1; 199/1; 307/1

C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 16

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin

F:16-77/Domain: immunoglobulin homology <IMM>

F:115-183/Domain: immunoglobulin homology <IMM>

F:220-288/Domain: immunoglobulin homology <IMM>

F:325-396/Domain: immunoglobulin homology <IMM>

F:123-75,122-181,221-286,332-394/Disulfide bonds: #status predicted

F:43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.5%; Score 363.5; DB 1; Length 423;

Best Local Similarity 29.8%; Pred. No. 2e-14; Indels 61; Gaps 17;

Matches 103; Conservative 68; Mismatches 114; Indels 61; Gaps 17;

```

Qy      109 CEVEDQKEVQD--LVFGLTAN--SDTHLQGSLLTLSPGSSPVSQCRSPRKNIQ 164
Db      109 CDPAFHSTIDLYCYTHINDVSWSLMDRETDL-----AQVILKE--EGKLAS 161
Qy      165 GSKTISVQLELDQSGTCTVVLQNKVEFKIDIVCPAPRPSCKDTHTCPELLGSPS 224
Db      162 TCSKLNTIEQOMWSESTFCRV--TSQGVLYLAHTRRCPDHEPR-----GAI 206
Qy      225 VFLPPPKDITLMTSRTEPVTCVVVDV--SHEDPEVKEN-----WYVDGVEVYNA 272
Db      207 TYLLPSPSLD-LYQNGAPKLTCLVVDLSEKVNNTVMQEKTSVSASQWY--TKHN- 261
Qy      273 KTKPREQNSTYRVVSVLYTLHODPLNGEKYCKVSKALPAPIEKTISYAK--GQPREP 331
Db      262 -----NATSTISILPVAAKDMINGYQCVVDPPDPKPRIVRSITILPQVSQSRAP 312
Qy      332 QVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENN--YKTPPVLDSDGS- 388
Db      313 EYVYFPPPEEE--SEKRTLTCLIQVFPEDISVQWLGSKLISNSQHSSTTP--LKSNGSN 370
Qy      389 --FPLYSKLTVDKSMQGNFSCSVNHEALHN--HYTKSLSPG 431
Db      371 QGFPLFSRLVAKTLMWRKQPTCOVHIALQKPRKLEKTIISTSLG 416

```

RESULT 66

A24976

Ig mu chain C region (allele b) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 21-Jan-2000

```

C:Accession: A24976
R:Schröter, P.H.; Quester, S.; Bothwell, A.
Nucleic Acids Res. 14, 2381-2389, 1986
A:Title: Alleotypic differences in murine mu-genes.
A:Reference number: A24976; MUID:86176735; PMID:3083402
A:Accession: A24976
A:Molecule type: mRNA
A:Residues: 1-455 <SCH>
A:Cross-references: GB:X03690; NID:952381; PID:CA27326.1; PID:952382
A:Experimental source: strain C57BL/6
A>Note: the authors translated the codon AAG for residue 65 as Leu
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:346-417/Domain: immunoglobulin homology <IMM>

```

Query Match 13.4%; Score 363; DB 2; Length 455;

Best Local Similarity 25.3%; Pred. No. 2.3e-14;

Matches 116; Conservative 84; Mismatches 149; Indels 110; Gaps 22;

```

Qy      37 VELTCTASQ--KKSIOFHMKNNOIKILQNGSFLTKGPKSLNDRASRLMDQGNF-- 92
Db      24 VAMGCLARDLPSTISTYNYQNNTVEYIOGIRTPPT-----LRTGKLYA 68
Qy      93 ---PLIKNLKIEDSTY-ICEVEDQKEVQLVFGTLTANSDDL-----LQGSLLTLE 144
Db      69 TSQVLSPKSILEGSDYLVCIKH-----YG-GKNKDLHVPAPVAKMPPNVVF 117
Qy      145 SP-----GSSP---SVQCR-----SPR-----GKNIOG----- 166
Db      118 VPPDGGSGAPRPSKILCATATNFTPRITVSWLKDGLVESGTTDPVTIENKSPQT 177
Qy      167 -----KTLVSQLELDQSGTCTVVLQNKVEFKIDIVCPAPRPSCKDTHTCPELLG 222
Db      178 YKVISTILFIEIDMLNANVTCTV--DHRGLTFLKAVSSICAAAPST-----DIL-- 225
Qy      223 PSVFLPPKPRDITLMSRTEPVTCVVVDVSHEDPEVKFNMYVDGVEVNAKTPREQYN 282
Db      226 -FTTIPSPAD-IFLSKSNLTLCLVSNLTYE-TLNISWASQGEPLKIKIMESHFN 281
Qy      283 STYVSVSVLYTLHODPLNGEKYCKVSKALPAPIEKTISYAKQPRE-----QVYTLR 337
Db      282 GTFSAKVAIVAVCEDMNNRKEFVCTVTRDLPSPQKFKISK---PNEVHGRPAVYLLP 337
Qy      338 PSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQ--PENNYKTPPVLD--SDGSFPLY 392
Db      338 PARQLMRSATYTCIVKGFSPADISVQWLGSKLISNSQHSSTTP--LKSNGSN 397
Qy      393 SKLTVDKSMQGNFSCSVNHEALHNHYTKSLSPG 431
Db      398 SILVTBEENWNGETVTCVVGHEALPHLVTERVTVDKSTG 436

```

RESULT 67

EMMS

Ig mu chain C region, secreted form - mouse

C:Species: Mus musculus (house mouse)

C:Date: 31-Oct-1980 #sequence_revision 31-Oct-1980 #text_change 16-Aug-1996

C:Accession: A02166; A26239; A26240; B02039

R:Kawakami, T.; Takahashi, N.; Honjo, T.

Nucleic Acids Res. 8, 3933-3945, 1980

A:Title: Complete nucleotide sequence of mouse immunoglobulin mu gene and comparison with

A:Reference number: A02166; MUID:81076590; PMID:6255422

A:Accession: A02166

A:Molecule type: DNA

A:Residues: 1-455 <KAW>

A:Cross-references: GB:J00443

A>Note: the sequence was determined from the germline gene

R:Goldberg, G.I.; Vanin, E.F.; Zrolka, A.M.; Blattner, F.R.

Gene 15, 33-42, 1981

A:Title: Sequence of the gene for the constant region of the mu chain of Balb/c mouse

A:Reference number: A26239; MUID:82051295; PMID:6795090

A:Accession: A26239

A:Molecule type: DNA

A:Residues: 1-455 <GOL>
 A:Note: the sequence was determined from the germ-line gene
 R:Auffray, C.; Rougeon, F.
 Gene 12, 77-86, 1980
 A>Title: Nucleotide sequence of a cloned cDNA corresponding to secreted mu chain of mouse
 A:Reference number: A26240; MUID:81165562; PMID:6260591
 A:Contents: TEPC183
 A:Accession: A26240
 A:Molecule type: mRNA
 A:Residues: 1-225, 'N', 227-257, 'S', 259-367, 'K', 369-455 <NUP>
 R:Kenly, M.; Sibley, C.; Fuhrman, J.; Schilling, J.; Hood, L.E.
 Proc. Natl. Acad. Sci. U.S.A. 76, 2932-2936, 1979
 A:Reference number: A26241; MUID:79223904; PMID:111247
 A:Contents: annotation; MOPC 104E
 A:Note: this sequence has been revised in reference A202039. Carbohydrate binding sites at
 R:Kenly, M.R.; Fuhrman, J.S.; Schilling, J.W.; Rogers, J.; Sibley, C.H.; Hood, L.E.
 Biochemistry 21, 5415-5424, 1982
 A>Title: Complete amino acid sequence of a mouse mu chain: homology among heavy chain cc
 A:Reference number: A202039; MUID:83075344; PMID:6816276
 A:Contents: MOPC 104E
 A:Accession: B02039
 A:Molecule type: protein
 A:Residues: 1-77, 'N', 79-100, 'Q', 102-225, 'N', 227-257, 'T', 259-367, 'K', 369-455 <KEH>
 C:Genetics:
 A:Introns: 1/1; 106/1; 219/1; 325/1
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
 hain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1a
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob
 F:21-91/Domain: immunoglobulin homology <IMM1>
 F:129-201/Domain: immunoglobulin homology <IMM2>
 F:239-307/Domain: immunoglobulin homology <IMM3>
 F:346-417/Domain: immunoglobulin homology <IMM4>
 F:436-455/Domain: carboxyl-terminal <CTRS>
 F:1/4/Disulfide bonds: interchain (to light chain) #status predicted
 F:128-89/Disulfide bonds: #status experimental
 F:46, 211, 243, 281, 442/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:136-199, 246-305, 353-415/Disulfide bonds: #status predicted
 F:216, 455/Disulfide bonds: interchain (to heavy chain) #status predicted
 F:293/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 13.4%; Score 362; DB 1; Length 455;
 Beeg Local Similarity 25.3%; Pred. No. 2,6e-14;
 Matches 116; Conservative 84; Mismatches 149; Indels 110; Gaps 22;

QY 37 VELTCTASQ--KKSIOFHKNKSNQIKILNQGSLFKGFSKLNDRSRRSLMDQNF-- 92
 DB 24 VAMGCLARDFLPSTISFTWNYONNTEVIGIRTPPT-----LRTGKXLA 68
 QY 93 ---PLIKLKIKIEDDTY-ICEVEDQKEVQLVGLTANSPTHL-----LOGQSLTLTLE 144
 DB 69 TSQVLLSPKSLIEGSEYLVCKIH-----YG-GKNRDLHVPAPVAEMNNVNVF 117
 QY 145 SPP-----GSSP---SVQCR---SPR-----GKNIOGQ----- 166
 DB 118 VPRRGFGSGPARPKSKLCEATNFTPKPTTWSMLKDGKLVESGFTTDPVTIENKSTPOT 177
 QY 167 ---KTLVSQLELDQSGTWTCTVLONQKVEFKIDIVCPAPRPSKCDKTHTCBELLGG 222
 DB 178 YKVIKTLTITSEIDMINLVYTCRV--DHRGLTFELKNVSTCAASPST-----DIL-- 225
 QY 223 PSVFLPPPKKDTLMISRPEVTCVVVDVSHEDPEVKFMYWYDGVYVNAKTKRPEQYN 282
 DB 226 --TFITPSFAD-IFLISKANLITCLVSNLATE-TLNTSWASQSGEPLETIKIKIMESHFN 281
 QY 283 STYRVSVTLVHOPMLNGKEYCKVSNKALPAPIETISKAKGPR-----POVYTLR 337
 DB 282 GTFSAKGVASVCEVDENNRKEFVCTYTHDLSPQKKRISK-----PNEVHKRPRAVYLLR 337
 QY 338 PSRDEL-T-KNQVSLTCLVKGFPYSDIAVWESNGQ--PENNYKTPPVLD--SDGSPFLY 392
 DB 338 PAREQLNLRSAATVCTLVGFGSPADISVQWLORGQLLPQEKVYTSAPMPEGPAPGFYFTH 397

QY 393 SKLTVDKSRMOQGNVFCSVNHEALHNHYTKSISLSPG 431
 DB 398 SILVTTEENWNSGETYTCVVGHEALPHLVTERTVDXKSG 436

RESULT 68
 136948
 Ig epsilon-chain - chimpanzee (fragment)
 C:Species: Pan troglodytes (chimpanzee)
 C:Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
 C:Accession: 136948
 R:Sakoyama, Y.; Hong, K.
 Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987
 A>Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and oranguta
 A:Reference number: 136948; MUID:87147196; PMID:3103123
 A:Accession: 136948
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-426 <RES>
 A:Cross-references: GB:M15398; NID:G176797; PID:NAA35416.1; PID:G176798
 C:Genetics:
 A:Introns: 103/1; 209/1; 317/1
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 F:336-405/Domain: immunoglobulin homology <IMM>

Query Match 13.4%; Score 361.5; DB 2; Length 426;
 Beeg Local Similarity 28.9%; Pred. No. 2,6e-14;
 Matches 102; Conservative 54; Mismatches 136; Indels 61; Gaps 10;

QY 90 GNFPLILNKLIKIEDDTYICEVEDQKEVQLVFGLTAN--SDTHLOGQSLTLTLESPP 147
 DB 123 GHFPPTQLL-----CLVSGYTFGTINITLDEGQVWDVDL----- 158
 QY 148 GSSPSVQCRSPRGKNIQSGKTLVSQLELDQSGTWTCTVLONQKVEKIDIVCPAPAP 207
 DB 159 ---STASATBGEGLASTQSELTLSQKMLSDRYTCQVTVQSGTFE----- 201
 QY 208 KSCDKTHCPQL-LGSPVFLPPPKKDTLMISRPEVTCVVVDVSHEDPEVKFMYWYDG 266
 DB 202 --DSTKCAQSNPRGSAVLSRSPFD-LTRKSPITTLCLVVDLAPSKGTNLTWSPAS 257
 QY 267 VEVNNAKTPREQYNSTYRVSVTLVHOPMLNGKEYCKVSNKALPAPIETISKAKG 326
 DB 258 GKPVNHSTRKEQKQNGNLITVSTLPVGTNRMBETVQCAVTHHDLRALVRSSTKSG 317
 QY 327 QPREQVYTL-----PSRDELTKNQVSLTCLVKGFPYSDIAVWESN--GQPENNYKT 379
 DB 318 PRAAPEVYAFATPREGSGRDKRT-----LACLIQFMPEDISVQWLHNEVOLPDARHSTT 372
 QY 380 PVLVDSGSPFLYSLKTLVDKSRMOQGNVFCSVNHEALHNHYT--OKSLSLSPG 431
 DB 373 QPHKTKGSGFVFSRLVETRAWEQKDEFICRAVHAASPSQTVQRTVSNVP 425

RESULT 69
 EHHU
 Ig epsilon chain C region - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Mar-1981 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999
 C:Accession: A22771; A23195; PH1214; A93491; A90824; A94418; B93933; S02438; A53116; C46
 R:Flanagan, J.G.; Rablitz, T.H.
 EMBO J. 1, 655-660, 1982
 A>Title: The sequence of a human immunoglobulin epsilon chain constant region gene
 A:Reference number: A22771; MUID:84236029; PMID:6234164
 A:Accession: A22771
 A:Molecule type: DNA
 A:Residues: 1-428 <FLA>
 A:Cross-references: GB:L00022; GB:V00555; NID:G185035
 R:Ueda, S.; Nakai, S.; Nishida, Y.; Hasejima, H.; Honjo, T.
 EMBO J. 1, 1539-1544, 1982
 A>Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseudog
 A:Reference number: A23195; MUID:84207910; PMID:6327276
 A:Accession: A23195

A:Molecule type: DNA
A:Residues: 2-428 <U>
A:Cross-references: GB:J00222; NID:g184755
R:Zhang, K.; Saxon, A.; Max, E.E.
J. Exp. Med. 176, 233-243, 1992
A:Title: Two unusual forms of human immunoglobulin E encoded by alternative RNA splicing
A:Reference number: PH1214; MUID:92308839; PMID:1613458
A:Accession: PH1214
A:Molecule type: DNA
A:Residues: 320-428 <ZHA>
A:Cross-references: EMBL:X63693; GB:S38668; NID:g32987
R:Seno, M.; Kurokawa, T.; Ono, Y.; Ono, H.; Sasada, R.; Igarashi, K.; Kikuchi, M.; Sugita, K.
Nucleic Acids Res. 11, 719-726, 1983
A:Title: Molecular cloning and nucleotide sequencing of human immunoglobulin epsilon chain
A:Reference number: A93491; MUID:83168897; PMID:6300763
A:Accession: A93491
A:Molecule type: mRNA
A:Residues: 1-428 <SEN>
A:Cross-references: GB:L00022; GB:V00555; NID:g185035
R:Max, E.E.; Battey, J.; Ney, R.; Kirsch, I.R.; Leder, P.
Cell 29, 691-699, 1982
A:Title: Duplication and deletion in the human immunoglobulin epsilon gene.
A:Reference number: A90824; MUID:83001945; PMID:6288268
A:Accession: A90824
A:Molecule type: DNA
A:Residues: 1-358, 'L', 360-428 <MAX>
A:Cross-references: GB:J00222; NID:g184755
A:Note: this sequence difference may be due to polymorphism
R:Benich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.
in Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.1-3
A:Reference number: A94418
A:Accession: A94418
A:Molecule type: protein
A:Residues: 'GMWL', 6, 'X', 8-16, 'B', 18-43, 'B', 45-52, 55-92, 95-97, 'B', 99-121, 'B', 123, 'L', 124
A:Experimental source: myeloma protein Nd
R:Kerten, J.H.; Molgaard, H.V.; Houghton, M.; Darbyshire, R.B.; Viney, J.; Bell, L.O.; G
Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982
A:Title: Cloning and sequence determination of the gene for the human immunoglobulin epsilon
A:Reference number: A93933; MUID:83065234; PMID:6815656
A:Accession: B93933
A:Molecule type: mRNA
A:Residues: 1-40:58-114:427-428 <KEN>
A:Cross-references: GB:L00022; NID:g185035
R:Ikeyama, S.
FEBS Lett. 224, 306-310, 1987
A:Title: Purification and characterization of a recombinant human IgE Fc-epsilon fragment
A:Reference number: S02438; MUID:88083554; PMID:3121387
A:Accession: S02438
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 98-352 <IKE>
R:Zhang, K.; Max, E.E.; Cheah, H.K.; Saxon, A.
J. Biol. Chem. 269, 456-462, 1994
A:Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produces
A:Reference number: A53116; MUID:94105254; PMID:8276835
A:Accession: A53116
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 320-428 <ZHA>
A:Experimental source: myeloma U266-derived cell line AF-10
A:Note: sequence extracted from NCBI Backbone (NCBIN:141701, NCBI:P:141702)
R:Hellman, L.
Eur. J. Immunol. 23, 159-167, 1993
A:Title: Characterization of four novel epsilon chain mRNA and a comparative analysis of
A:Reference number: A46536; MUID:93122085; PMID:8419166
A:Accession: C46536
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 382-426 <HEL>
A:Cross-references: GB:S55273; NID:g263166; PIDN:ABB24857.1; PID:g263167
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI Backbone (NCBI:P:125297)
A:Accession: D46536

```

A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 382-391 <HE2>
A:Cross-references: GB:S55276; NID:G263168; PIDN:AAB24858.1; PID:G263169
A:Experimental source: B cell myeloma U-266
A>Note: sequence extracted from NCBI backbone (NCBIP:125299)
A:Accession: A46516
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 401-428 <HE3>
A:Cross-references: GB:S53497; NID:G263162; PIDN:AAB24855.1; PID:G263163
A:Experimental source: B cell myeloma U-266
A>Note: sequence extracted from NCBI backbone (NCBIP:123483)
C:Genetics:
A:Gene: GDB:IGHB
A:Cross-references: GDB:119335; OMIM:147180
A:Map position: 14q32.33-14q32.33
A:Introns: 1/1; 104/1; 211/1; 319/1
C:Complex: An immunoglobulin heterodimer subunit consists of two identical light (kap  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterodimer; immunoglobu  
F:12-87/Domain: immunoglobulin homology <IM1>
F:128-195/Domain: immunoglobulin homology <IM2>
F:232-301/Domain: immunoglobulin homology <IM3>
F:338-407/Domain: immunoglobulin homology <IM4>
F:14/Disulfide bonds: interchain (to light chain) #status predicted
F:15-105-29-85-135-293-299-345-405/Disulfide bonds: #status predicted
F:21-49-99-146-252-275/Binding site: carbohydrate (Asn) #status experimental
F:121-209/Disulfide bonds: interchain (to heavy chain) #status predicted

Query Match 13.4%; Score 361.5; DB 1; Length 428;
Best Local Similarity 28.9%; Pred. No. 2,6e-14;
Matches 102; Conservative 53; Mismatches 137; Indels 61; Gaps 10;

Qy 90 GNFPLIKNLKIEDSDTYICEVEDQKEEVOGLVFGILTAN--SDTHLQGGSLTLTLESPP 147
Db 125 GHFPPRTQL-----CLVSGTTPGTINTLTWLDGQVDDVL----- 160
Qy 148 GSSPVEQCRSPRGKNIQGGKTLVSQLELDSDSTWTCVTVAONQKVEFKIDIVPCPAEP 207
Db 161 ----STASTQGGELASTQSELTLSSQKMLSDTYTCQVYVYQGHTE----- 203
Qy 208 KSCDKHTTCPEL-LGGPSVFLPPPKDKDTLMTSRTEBVCVVVDVSHEDPEVKENMYVDG 266
Db 204 ---DSTKKCADSNPRGVSAYLSRSPSPD-LFIRKSPITCLVVDLAPSKGTNLTWSRAS 259
Qy 267 VEVNHAKTKRREOVNSTRYVSVLTVLHODMLNGEKYKKKNSKALPAPREKITSKRG 326
Db 260 GKPVNHSTRKEKQRNGTLTVSTLPGTRDWTGEGTYQCRVTHPHLPALMRSTTKSG 319
Qy 327 QPREQVYTL-----PPSRDELTKNQVSLTCLAKGFPSPDIAYVESGN--GQPNYKTT 379
Db 320 PRAAEVVAFAFPMPGSDKRT----LAQLQNPEPDISVQMLNNEVQLPARKHSTT 374
Qy 380 PPVLSDSDSFFPLYSKLTVDKSRWQGNVFSQSVMEALNHYT-QKSLSPG 431
Db 375 QPKTKGSGFVFYSRLLEVTRAEMEQKDFICRAVHEAASPQTVQRAVSVVPG 427

RESULT 70
S00390
I9 gamma chain (clone 36) - chicken (fragment)
N:Alternate names: Ig nu chain
C:Species: Gallus gallus (chicken)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
C:Accession: S00390
R:Parvart, R.; Avital, A.; Lentner, F.; Ziv, E.; Tel-Or, S.; Burestein, Y.; Schechter, I.  
EMBO J. 7, 739-744, 1988
A:Title: Chicken immunoglobulin gamma-heavy chains: limited VH gene repertoire, combinat  
A:Reference number: S00390; MUID:88283642; PMID:3135182
A:Accession: S00390
A:Molecule type: mRNA

```

A:Residues: 1-504 <PAR>
 A:Cross-references: EMBL:X07174
 A:Note: this sequence was determined from the differentiated gene
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: immunoglobulin

Query Match 13.4%; Score 361; DB 2; Length 504;
 Best Local Similarity 28.3%; Pred. No. 3.4e-14;
 Matches 95; Conservative 70; Mismatches 129; Indels 42; Gaps 14;

QY 114 QKEEVL--LVFGLT-ANDTH-LLOG-OSLTLESPPGSSPVOCRRPKNIQGGKT 168
 DB 192 QSEVELLCVYGFSPASAEVEMLYDVGGLVVASQSPAVRSGSTYLSR----- 242
 QY 169 LVSQLELDOSGWTCTVQONOKKVEFKIDIVCPAPRPEKSCDKHTCCPELLGGSPVLF 228
 DB 243 VAVSGTDMREGKSYSCRVHPATNTVYEHVKGCP-----DGAQSCSPI---QLVNI 291
 QY 229 PRPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPR--EEOYNSTYR 286
 DB 292 PPSGPE-LYISDAKRLKCLVNLPR-SDSSLVTWIRE--KSGNLPRDPVNLQEHFNGTYS 347
 QY 287 VVSULTVLHQMUNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKN 346
 DB 348 ASSAVPVSTQDWLSGERFCTVQHEELPLPLSKSVYRNTGPTTPPLIYFPADHPELISLS 407
 QY 347 QVSLTCLVGYFSPDIAVEM--ESNQGPENNYKTPPVLD-----SDG-SFLYSLK 395
 DB 408 RYTLSCLVAGFPRDRIEIMLDHRAVPRTEFVYTAVALPEERTANGAGDGDTPFYVSKM 467
 QY 396 TVDKSRMQQGNFSCSVMEALHNHYTQKSLSLSPG 431
 DB 468 SVETAKMNGVTVPACMAVHAELPMRSPQRTLOKQG 503

RESULT 71

MHRB

Ig mu chain C region, secreted form - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 23-Aug-1997
 C:Accession: A02164
 R:Bernstein, K.E.; Alexander, C.B.; Reddy, E.P.; Mage, R.G.
 J:Immunol. 132, 490-495, 1984
 A:Title: Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain of V-Haz 8
 A:Reference number: A02164; MUID:84088930; PMID:6418603
 A:Accession: A2 A1locType
 A:Accession: A02164
 A:Residues: 1-458 <BER>
 A:Molecule type: mRNA
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobulin
 F:21-92/Domain: immunoglobulin homology <IMM1>
 F:130-202/Domain: immunoglobulin homology <IMM2>
 F:242-310/Domain: immunoglobulin homology <IMM3>
 F:349-420/Domain: immunoglobulin homology <IMM4>
 F:14/Diulfide bonds: interchain (to light chain) #status predicted
 F:28-90,137-200,249-308,356-418/Diulfide bonds: #status predicted
 F:46,114,212,261,277,284,445/Binding site: carbohydrate (Asn) (covalent) #status predict
 F:219,457/Diulfide bonds: interchain (to heavy chain) #status predicted
 F:296/Diulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 13.3%; Score 360; DB 1; Length 458;
 Best Local Similarity 23.8%; Pred. No. 3.5e-14;
 Matches 118; Conservative 76; Mismatches 145; Indels 156; Gaps 22;

QY 21 AATGCKVVLGKKGDVVELTCTASQ--KSIQFHW--KSNQIKILGNGSLTKGPKSL 76
 DB 17 ALTDGKLVAMG-----CLARDLPSSVTFSSFKNSBI----- 50
 QY 77 NDRAUSRSRLMDQGNFLIIKILK-----IEDDTY-ICEVEDQKEVQL 121

DB 51 -----SRTV---RTFVVKRGDKYMATSOVLVPSKDLQGTBEYLCKVQHSNMRDLR 102
 QY 122 VFGLTANDTHLLOQSLTLTLESPGSSPVQGR-----SPR-- 159
 DB 103 V-----SEPVSELPNNVSVFIPPDSSGSGSTRKSRLLCOATGFSFKOI 147

QY 160 -----GKNIQGG-----KTLVSQLELDOSGTW-----TCTV-----L 187
 DB 148 SVSMRLDQKVBSSGLTPVVEAETKAGAPATPFSISMLTTESDVLSSGLYTCRDHNGI 207
 QY 188 QNOKVEFKIDIVCPAPRPEKSCDKHTCCPELLGSPVFLPPPKDTLMISRTPEVTCV 247
 DB 208 FFDKKNVSSSECSSTPSP-----GIQVFIAPSPADT-FLSKSARLICT 250
 QY 248 VVDVSHEDPEVKFNWYVDGVEVHNAKTKPR--EEOYNSTYRVSULTVLHQMUNGK 302
 DB 251 VDTLTYTG-SLNISSV-----ASHNKGALDTHNNTESHPPNATFSGMGAASVCAEDWESGE 304
 QY 303 EYKCVSNKALPAPIEKTISKAKGQPRE-PQVYTLPPSRDELTK-KNQVSLTCLVGYFYS 360
 DB 305 QPTCTVTHADLPFLPKHTISKREYAKIPRAVYVLPARQGLVRESATVTCVYGFSPA 364
 QY 361 DIAVESNQGPR--ENNYKTPPVLD--SDGSFFLYSKLYDKSRMQQGNFSCSVMEHA 416
 DB 365 DVFVQMQGQPLSDSKYVTSAPADEPOAPGLYFHTSLTYTEEDMNGGEFTCVVGHBA 424
 QY 417 LHNHYTQKSLSLSPG 431
 DB 425 LPHMYTERTVDKSTG 439

RESULT 72

160082

CD4 receptor - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 23-Jul-1999
 C:Accession: I60082
 R:Zverev, V.V.; Sidorov, A.V.; Nedospasov, S.A.; Malinbova, V.V.; Udalova, I.A.; Andzha Vopr. Virolog. 40, 100-102, 1995
 A:Title: [Nucleotide sequence of two exons of the human T-lymphocyte CD4 receptor gene].
 A:Reference number: I60082; MUID:95407135; PMID:7676667
 A:Accession: I60082
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-71 <RES>
 A:Cross-references: GB:S79267; NID:G1086922; PIDN:AAAB35273.1; PID:G1086923
 A:Genetics:
 A:Introns: 17/1
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

Query Match 13.2%; Score 357; DB 2; Length 71;
 Best Local Similarity 98.6%; Pred. No. 5.7e-15;
 Matches 70; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFPHLLVLTALPPAATGKNKVLGKKGDVVELTCTASQKSIQFHWKSNQIK 60
 DB 1 MNRGVPFPHLLVLTALPPAATGKNKVLGKKGDVVELTCTASQKSIQFHWKSNQIK 60
 QY 61 ILGNGSFLTK 71
 DB 61 ILGNGSFLTK 71

RESULT 73

S03961

Ig mu chain C region - house shrew (fragment)
 C:Species: Suncus murinus (house shrew)
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jan-2000
 C:Accession: S03961
 R:Ishiguro, H.; Ichihara, Y.; Namikawa, T.; Nagatsu, T.; Kuroawa, Y.
 FBS Lett. 247, 131-132, 1989
 A:Title: Nucleotide sequence of Suncus murinus immunoglobulin mu gene and comparison wit
 A:Reference number: S03961; MUID:89232144; PMID:2497033

Db 277 SHPNGTFAIGBANVCVEDMDSGKEFVCTVTRHDLPSPOKFKISPRGMNKTTPAVYQCP 336
 Qy 338 PSRDEL-TKNQVSLTCLVGVFPSPDIWESNCGP--ENNKTTPVLDSDGS--FLY 392
 Db 337 LAREQILLESATVTLVGVSPADLVGMORQGLSDQKVTSAFMKEPAPHLVFTH 396
 Qy 393 SKLTVDKSRWQGVNFCSCVMEALHNHYTKSLSPG 431
 Db 397 SVLTJTEEBMNGERTYTCVGHGHALPHMVTERTVDRSTG 435

RESULT 76

C31933

Ig mu chain C region - African clawed frog (fragment)

C Species: Xenopus laevis (African clawed frog)

C Date: 31-Mar-1990 #sequence_revision 30-Jun-1992 #text_change 20-Sep-1999

C Accession: C31933

R: Schwager, J.; Mikoyak, C.A.; Steiner, L.A.

Proc. Natl. Acad. Sci. U.S.A. 85, 2245-2249, 1988

A: Title: Amino acid sequence of heavy chain from Xenopus laevis Igm deduced from cDNA se

A: Reference number: A94192; PMID:88176921; PMID:2451244

A: Accession: C31933

A: Molecule type: mRNA

A: Residues: 1-453 <SCH>

C: Superfamily: Immunoglobulin C region; immunoglobulin homology

C: Keywords: immunoglobulin

Query Match 12.8%; Score 345.5; DB 2; Length 453;
 Best Local Similarity 25.8%; Pred. No. 2.5e-13;
 Matches 119; Conservative 77; Mismatches 189; Indels 77; Gaps 18;

Qy 31 GKKGPVELTCTASQ--KKSIOFMKNSQIKLNGQSF---LTGKPSKLNDRADSR 84
 Db 15 GSGMDPVTIGCLAKDFLPETISTWGDKNASYSIGLKYPVMOSSGYTSSQVNVAS 74
 Qy 85 SLMDQGNFLIINKLKIETSDTYICEVEDQKEVQ-----LTV 122
 Db 75 AAMD--NIQFPCNAGHLDT---IKSVELKDPVKVEKPVVSIHPSKDALALNESLFI 129
 Qy 123 FGLTAN-SDTHLLQGQSLTLTLESPGSSPVQCRSP--RGNKIQGKTLVSQLELD 178
 Db 130 VCLATFTPTHTI-----VIKMLKNGQGTGVRVEBPDKRGVATSYLSITRKEMLD 184
 Qy 179 SGTWCTVYQNGKVEFKIDYPCAPBEKSCDKTKTCGELLGSP--VFLPRPKQTL 236
 Db 185 DTLVSCV-----EHAESGSLQEKMSKSLMCTPTPTPSIQVITTPS-LESI 232
 Qy 237 MISRPPEVTCVVVDVSHEDPEVKFNNY--VDGVEVHNAKTKRPEBOYNS--TYRVVSVLTV 293
 Db 233 PEKKATITLCLVSNMANSDELRSISYFKKSGTOEILKTELDALYNDKRTYSVKQTTV 292
 Qy 294 LHDWLNGKYEYCKVSNKALPAPIETKISKAGQPREPOVYTLPSRDELTKNQ-VSLTC 352
 Db 293 CADENNNDK-FVCKVEHTELASWKEVFLPEKEGYVTPVYVFPPELELSKRETTALTC 351
 Qy 353 LYKGFPSPIAYEM--ESNGQPENNYKTP-----PVLDSDGSFPLYSKLVNDSRWQ 404
 Db 352 LVKGFPSPEIFVYKMLHKNVAVPKQNYINTSINDELLPKQKSGFELYSLHTIDIKMDWA 411
 Qy 405 GNVSFCSVMEALHNHYTKSLSPG-----LQDETGC 438
 Db 412 GDSFSCVGHESLPLOLTORSIDKSGKFTNNVNSLVLSDTG 453

RESULT 77

B30503

Ig gamma-2a chain C region (B5.7A12) - mouse (fragment)

C Species: Mus musculus (house mouse)

C Date: 31-Mar-1989 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999

C Accession: B30503

R: Gilmore, G.L.; Bard, J.A.; Birnstein, B.K.

J. Immunol. 141, 1754-1761, 1988

A: Title: DNA rearrangements affecting both variable and constant regions of Ig H chain g

A: Reference number: A30503; PMID:88315788; PMID:2842402
 A: Accession: B30503
 A: Molecule type: mRNA
 A: Residues: 1-112 <GIL>
 A: Cross-references: GB:M21925
 A: Experimental source: myeloma cell line MPC11
 A: Note: the authors translated the codon GAG for residue 11 as Ser
 C: Genes: A: Introns: 100/3
 C: Superfamily: immunoglobulin C region; immunoglobulin homology
 C: Keywords: immunoglobulin
 F: 1-70/Domain: immunoglobulin homology <IMM>

Query Match 12.3%; Score 331.5; DB 2; Length 112;
 Best Local Similarity 50.0%; Pred. No. 3.2e-13;
 Matches 66; Conservative 17; Mismatches 22; Indels 27; Gaps 2;

Qy 239 SRTPEVTCVVVDVSHEDPEVKFNNYVDGVEVHNAKTKRPEBOYNSYRVSVLTVLHDQ 298
 Db 1 SLTPKVTCTVVDVSDDDVDQISWPNVNEVHTAQOTHRREDYNTIRVSVTLPIQHDW 60
 Qy 299 LNKGEYCKVSNKALPAPIETKISKAGQPREPOVYTLPSRDELTKNQVSLTCLVKERY 358
 Db 61 MSGKEPKCKVNNKDLPAPIERTISKPE-----SCSLTAWGLGW- 100
 Qy 359 PSDIWEWESNG 370
 Db 101 -----WTSNG 105

RESULT 78

I50731

Ig heavy chain - nurse shark

C Species: Ginglymostoma cirratum (nurse shark)

C Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C Accession: I50731

R: Vazquez, M.; Mizuki, N.; Flajnik, M.F.; McKinney, E.C.; Kasahara, M.

A: Title: Nucleotide sequence of a nurse shark immunoglobulin heavy chain cDNA clone.

A: Reference number: I50731; PMID:92357056; PMID:1495502

A: Accession: I50731

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: mRNA

A: Residues: 1-577 <VAZ>

A: Cross-references: GB:M2851; NID:9213264; PIDN:AAA50817.1; PID:9213265

C: Genes: IGH

A: Gene: IGH

C: Superfamily: immunoglobulin C region; immunoglobulin homology

F: 469-539/Domain: immunoglobulin homology <IMM>

Query Match 12.0%; Score 325.5; DB 2; Length 577;
 Best Local Similarity 24.3%; Pred. No. 5.1e-12;
 Matches 142; Conservative 83; Mismatches 165; Indels 195; Gaps 28;

Qy 10 LLLVQLALPAPATGKNV-----LGKKGPVELTCTASQKSIQFMKNSQ--IKTL 62
 Db 6 IFLSLALLALPCVQSEITLIDPEAFTHPGGSLTLC---KTFQNVGSSSMYMKOV 60
 Qy 63 GNOG-----SFLTGKPSKLNDRADSRSLMDQGN--FPLIKNLKIETSDTYICEV 111
 Db 61 PGQGLEWIVYVYSSSMNYPAPAIKDRFTAQ---DTSNNIFALNRSVKIDTAIYCC- 115
 Qy 112 EDQKEEVOLLVGLTANSDDLLOGQSILTLTESPPGSSPV-----QCRS----- 157
 Db 116 -----TRRMGYEVLGSHGYWGQGTWYVTTATP--SSPTLYGLVSSCOQNDIGSVLY 168
 Qy 158 -----PRGKIQGKTLVSQ----- 174
 Db 169 GCLAMDYPPDVASVTWKHGQLITTVGTVPVSRNKKGYTLLS-SQLALIESDAECDOIS 227
 Qy 175 -ELQDSG-----TWCT 185
 Db 228 CEVRHSGSDSKTGMCPDGFPTALLTVSSSEIESRKALIVCSISDFHSKISVLTW--- 284

```

QY 186 VLONOKYVEFKIDIVP-CPA-----PEPKSCDK-THTC-----PEL-----LGG 222
DB 285 -LKNGRSVDSGIFTSPVCEANGNFSVTSLRLRPVAMEMWDBRAVYTCQVAKKEVIOQSWNTG 343
QY 223 PSVF-----LFPKPKDITMISRTPEVTCVVVDVSHEDPEVKENNYVDGVEVHNAKT 274
DB 344 PVSSECHGYTAKILPP-PVEQVLEATVTITCV---VSNLSGQVFTWLOD-----EKT 393
QY 275 KPREQVNS---TYRVSVTLVTHODMNLGKCKVKCNKALPAIEKTIKAK-GOPRE 330
DB 334 LKSELAHNSGHSDAISKLDISTAMISEVFEVCVNHQYLPPLRDSIHKERLENPLE 453
QY 331 POVYTLPPSRDELTKNQ-VSLTCLVKGFYPSDIAVEWESNGQPEN--NYKTPPYLSDG 387
DB 454 PVSVTLPTTEELSNQRFSLTCLVARGFRPREIFPKMTTNNKPVNPSVYKNTVTAESDN 513
QY 388 -SFFLYSLITVDKSNWQGNVFCSCVMEBALHNTQSLSLSPG 431
DB 514 TSFFLYSLITIAERWASGASYSYCVGHEAIPLKTIINRTVDKSSG 558

```

RESULT 79

```

S31436
Ig upsilon chain - axolotl (fragment)
C:Species: Ambystoma mexicanum (axolotl)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999
C:Accession: S31436
R:Feilich, J.S.; Miles, M.V.; Schwager, J.; Charlemagne, J.
Submitted to the EMBL Data Library, November 1992
A:Description: cDNA sequence of Ambystoma mexicanum upsilon heavy Ig chain.
A:Reference number: S31436
A:Accession: S31436
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-433 <FEU>
A:Cross-references: EMBL:X69492; NID:G62420; PID:G62421
C:Superfamily: Immunoglobulin C region; Immunoglobulin homology

```

```

Query Match 11.9%; Score 321.5; DB 2; Length 433;
Best Local Similarity 24.2%; Pred. No. 6.3e-12;
Matches 111; Conservative 65; Mismatches 146; Indels 137; Gaps 20;

QY 28 VTLGKKGDTVEL-----TCTASQKKSIOFMKNSNQIKILGNSFLTKGPS 74
DB 28 LVTGLGPRVEVSWKAGSTMPGKTPPAQQAVAL--STSSQRIAYE----- 76
QY 75 KLNDRAADRSRLMDQGNFPLIINKLKIEDSDTYI-----CEVEDQKEEYQL----- 120
DB 77 -----WGTNSYSCSVKH--KPTSTEIHKITTSABCKKATKSPVQVQLQSSCA 121
QY 121 -----LVFGLT-----ANSDTHLQGSLSLTLESPPGSSPSVQCRRSP 159
DB 122 DTDGKSGIELVCLISGVTPDNIQVRLVNDKAPRIQGGT-----SPR 163
QY 160 GKNIGG---GKTLVSQLELDSDGTTCTVLONOKYVEFKIDIVPCAPAPKPSCKDTHT 215
DB 164 QKDGGTSTTSQINVTSDMASGDXYTC-----KVEH-----PATSSRAEDITHN 209
QY 216 CPELLIG--GPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENNYVDGVEVHNAK 273
DB 210 CADSGTPYQKVFLLAPKARD-LYIANQPVYICKITKKNNSP-SLSVTM-----KRR 259
QY 274 TKPRE-----EGY---NSTYRVSVTLVTHODMNLGKCKVKCNKALPAIEKTIKAK 325
DB 260 EGPEPAAVISSEQYIDSDGFTAMSLNTKKEWEGDFTCKVKHFDLPFLPSVSKSP 319
QY 326 GQPREPOVYTLPPSRDELTK-NQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTPPV 382
DB 320 GRSFAPMTYVFAFHMEELANDVFLSLTCLVKSFSFDDIYIQMKQCKSVIPSDKYVSMER 379
QY 383 LDSD---GSFFLYSLITVDKSNWQGNVFCSCVMEHAL 417

```

DB 380 QEAGTAGLGTFFSYSMULTIQKSDMDKRETFVAAHSAV 418

RESULT 80

```

A39016
T-cell surface glycoprotein CD7 precursor - human
N:Alternate names: T-cell leukemia antigen
C:Species: Homo sapiens (man)
C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 05-Nov-1999
C:Accession: A39016; S03520
R:Schenberg, L.E.; Fleener, D.E.; Kurtzberg, J.; Haynes, B.F.; Kaufman, R.E.
Proc. Natl. Acad. Sci. U.S.A. 88, 603-607, 1991
A:Title: Isolation and characterization of the genomic human CD7 gene: structural simila
A:Reference number: A39016; MUID:91110576; PMID:1703303
A:Accession: A39016
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-240 <SCH>
A:Cross-references: GB:M37271; NID:g180163; PIDN:AAA51953.1; PID:g180164
R:Aruffo, A.; Seed, B.
EMBO J. 6, 3313-3316, 1987
A:Title: Molecular cloning of two CD7 (T-cell leukemia antigen) cDNAs by a COS cell expr
A:Reference number: S03520; MUID:88111517; PMID:3501369
A:Accession: S03520
A:Molecule type: mRNA
A:Residues: 1-240 <ARU>
A:Cross-references: EMBL:X06180; NID:g29819; PIDN:CAA29546.1; PID:g732757
C:Genetics:
A:Gene: GDB:CD7
A:Cross-references: GDB:119770; OMIM:186820
A:Map position: 17q25.2-17q25.3
A:Introns: 28/1
C:Keywords: glycoprotein; membrane protein; surface antigen; tandem repeat
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-240/Product: T-cell surface glycoprotein CD7 #status predicted <MAT>
F:145-180/Region: 9-residue repeats (X-P-P-X-A-S-A-L-P)

```

```

Query Match 11.7%; Score 316.5; DB 2; Length 240;
Best Local Similarity 41.7%; Pred. No. 6.2e-12;
Matches 90; Conservative 23; Mismatches 54; Indels 49; Gaps 11;

```

```

QY 321 ISKAKGP-----REPQVYTLPPSRDELTKNOVSLTCL---VKGFY-----PS 360
DB 14 LALARGPLGALAAQEVQSGPHCTTVPGA-----SVNITCSISGIRGIYLRQLGPQ 67
QY 361 DIAWESNQGPNENYTTPEVLDSDGSEFLYSKLTVDKSRWQGN--VSCSVMHBAH 418
DB 68 DI-LYEDGVVPTTDRFRGRIT-DPSGS---QDNLTITMRLLQSDGTTCQATIEV-- 120
QY 419 NHYTQSLSLSPGLQLD--ETCAEAQDGEILDGLMTTPPAPASALPAPPTGSALPDPTAS 476
DB 121 NMYGSGTLVLTVEQSGGMRCSDA-----PPAPASALPAPPTGSALPDPTAS 168
QY 477 ALPPAPASALPALAVISFLIGIGIGACVLART 512
DB 169 ALPPAPASALPALAVISFLIGIGIGACVLARTQ 204

```

RESULT 81

```

HVRK2
Ig mu chain C region (clone 12022) - horn shark (fragment)
C:Species: Heterodontus francisci (horn shark)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: S00980
R:Kokub, F.; Hinds, K.; Litman, R.; Shambloct, M.J.; Litman, G.W.
EMBO J. 7, 1979-1988, 1988
A:Title: Complete structure and organization of immunoglobulin heavy chain constant reg
A:Reference number: S00980; MUID:88328985; PMID:3138109
A:Accession: S00980
A:Molecule type: mRNA
A:Residues: 1-438 <KOK>
A:Cross-references: EMBL:X07784; NID:g63963; PIDN:CAA30617.1; PID:g63964
A>Note: the sequence was determined from the differentiated gene

```

C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1a C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin
F.1-438/Domain: C region <CRE>
F.20-87/Domain: immunoglobulin homology <IM1>
F.123-190/Domain: immunoglobulin homology <IM2>
F.228-291/Domain: immunoglobulin homology <IM3>
F.330-400/Domain: immunoglobulin homology <IM4>
F.166,200,245,275,374,411,415,423/Binding site: carbohydrate (asn) (covalent) #status pr

Query Match 11.5%; Score 312; DB 1; Length 438;
Best Local Similarity 25.4%; Pred. No. 2.3e-11;

Matches 113; Conservative 67; Mismatches 155; Indels 110; Gaps 17;

QY 23 TGGNKV---VLGGKDTVELTCTASOKSIQFMKNSNQIKLGNOSFLTGPSKLTNR 79
DB 49 TTGLKIPSVLNKKGT-----YTRSSQLTITSEF-----VGSXKYCE 86
QY 80 ADSRSRLMDQGNPPLIKNLKIEDSDTYICEVEDQKEEYQLLVFGLTANSQTHL----- 133
DB 87 VARGESLW-----IKELDCKGDIIVPTVILQSSSEITSRRA 126
QY 134 -----LQGSFLT-----LESPPGSSPSVQCRSPRGKNIQSGKTLSSOLELQD 178
DB 127 TVLCSTIDPHPSITVSWLKDGQPMDSGFVTSPCTCVNG-----NFSATRLTVPAQWFS 182
QY 179 SGTWTCTVQLQNKKEFKI--DIVCPADPEKSCDKTTCPELGGPSVFLPPPKQDTL 236
DB 183 NTVTYTCQVAHOETQSRNITGSOVPCS-----IGDPVTKLPPSPTEQVL 226
QY 237 MSRPPEVTCVVDVSHEDPEVKFMNYVDGVEVHNAKTPREQY-----NSTRYVSVTL 291
DB 227 -LEATVTLTLCV---VSNADYGVNWSW-----TOEKPLSEIAVQGEQSDSYLSTV 274
QY 292 TVLHQMNLGKEXKCKVSKNALPAPIEKTISKAKGP-REPOV-YTLPPSRDELTKNQVS 349
DB 275 NSTQMLSGAEFYCVVSHQDLPTPLRASIHKEEVQDLREPFVSVALPPAEVDVSNQRFIS 334
QY 350 LTCLVKGFPYSDIAVWESNGQEN--NYKTPPVLDSPG-SFFLYSKLTVDKSRWQGN 406
DB 335 LTCLVKGFPYSDIAVWESNGQEN--NYKTPPVLDSPG-SFFLYSKLTVDKSRWQGN 394
QY 407 VFGSCVMEHMLNHNQYKSLSPG 431
DB 395 SYSCVVGHEALPILKIIRTNKSSG 419

RESULT 82

150830 Ig mu chain - Lepidosteus osseus (fragment)

C:Species: Lepidosteus osseus

C>Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 03-Nov-2000

C:Accession: 150830

Dev. Comp. Immunol. 19, 153-164, 1995

A>Title: cDNA sequences and organization of Igm heavy chain genes in two holostean fish.

A:Reference number: 150006; MUID:96039719; PMID:7556802

A:Accession: 150830

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-474 <HL>

A:Cross-references: EMBL:U12455; NID:9529949; PIDN:AAC59688.1; PID:9529950

A:Gene: Igm

C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 11.4%; Score 308.5; DB 2; Length 474;

Best Local Similarity 24.2%; Pred. No. 4.2e-11;

Matches 116; Conservative 62; Mismatches 171; Indels 131; Gaps 20;

QY 31 GKKGDTVELTCTAS--QKSIQFMKNSNQIKLGNOSFLTGPSKLTNRADRSRLMD 88
DB 31 GKKGDTVELTCTAS--QKSIQFMKNSNQIKLGNOSFLTGPSKLTNRADRSRLMD 88

DB 50 GTSGDVALGCLATGFLPDLSTLFKMTDSTDKEL-----TPRKRPSVLNGETYSSTS--- 101
QY 89 QGNFPLIKNLKIEDSD-----TYICEVEDQKEEYQLLVFGLTANSQTHLQGSFLTTL 143
DB 102 -----QSLPSPDMNSGKAFCEKHPQGDVYL-----HL----- 132

QY 144 ESPPSSPSVQCRSPRGKNIQSGKTLSSOLELQDSGTWCTVLO-NOKKVEFK----- 196

DB 133 -TPPVPASVLLNPF-----SLEBFAQNTATLVCVRSQGFSPRTHEFKMRRGNT 180

QY 197 ----IDIVPCPAPPK-----SCDKTH-----TC 216

DB 181 KLDEGVNTIPAVVDKTKYASSLTVEKQWKSAPACEVHTGSLVKNITYSRRC 240

QY 217 PELGGPSVFLPPPKQDTMISRTPEVTC---VVVDVSHEDPEVKFMNYVDGVEVHNAK 273

DB 241 QETV---KVLEPPTNEQ--EVKKTATLTCRIALVST---DVSMTSSGGKFL--NA 290

QY 274 TKPREQVNSTYRVSVLTVLHQMNLGKEXKCKVSN-KALPAPIEKTISKAKGPREFQ 332

DB 291 GAPERGHGCKIVANRSVLEKRTGTEYKCIYSHDSFPPTLTTKYKQDIATKIRS 350

QY 333 VYTLPPSDE--LYKNQVSLTCLVKGFPYSDIAVW--ESNGQENNYKTPPVLDSPG 388

DB 351 VFLAPSTEENSTRDEVTLTCEVDFSPKDIYISMLQDSIVDKTHVTLVDLPSHDGA 410

QY 369 FFLYKLTVDKSRWQGNVFGSVHMA-----LHNHYTKSLSLSPGLDFTC 438

DB 411 YSVSKYTISSSDMNSGTMYSQVHETAPLPVSVITRTDSTGKATLVNFTLPTDC 470

RESULT 83

S21461 T-cell surface glycoprotein CD4 (allele 1) - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 21-Jan-2000

C:Accession: 147131; S21461

J. Immunol. 151, 1365-1370, 1993

A>Title: Extensive allelic polymorphism in the CD4-like region of the miniature swine (C. guinea)

A:Reference number: 147131; MUID:93329116; PMID:8335933

A:Accession: 147131

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-99 <GU2>

A:Cross-references: EMBL:X65629; NID:91928; PIDN:CAA6563.1; PID:9388232

C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C:Keywords: glycoprotein; T-cell

F.3-81/Domain: immunoglobulin homology <IM4>

Query Match 11.3%; Score 305.5; DB 2; Length 99;

Best Local Similarity 60.2%; Pred. No. 9.7e-12;

Matches 59; Conservative 16; Mismatches 22; Indels 1; Gaps 1;

QY 32 KKGDTVELTCTASOKSIQFMKNSNQIKLGNOSFL-TKQSPKLTNRADRSRLMDPG 90

DB 1 KKGDTVELTCTASOKSIQFMKNSNQIKLGNOSFL-TKQSPKLTNRADRSRLMDPG 60

QY 91 NFPLIKNLKIIEDSDTYICEVEDQKEEYQLLVFGLTAN 128

DB 61 NFPLIKNLKIIEDSDTYICEVEDQKEEYQLLVFGLTAN 98

RESULT 84

HYRKKO

Ig mu chain C region, membrane-bound (clone 3050) - horn shark

C:Species: Heterodontus francisci (horn shark)

C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Aug-1996

C:Accession: S01854; C32716; A46530

R:Kokubu, F.; Hinder, K.; Litman, R.; Shambloot, M.J.; Litman, G.W.

EMBO J. 7, 1979-1988, 1988

A>Title: Complete structure and organization of immunoglobulin heavy chain constant region

A:Reference number: S00980; MUID:88328985; PMID:3138109

```

A:Accession: S01854
A:Molecule type: DNA
A:Residues: 1-461 <KOK>
A:Cross-references: EMBL:X07781
A>Note: the sequence was determined from the germ-line gene
R:Kokubu, F.; Hinde, K.; Litman, R.; Shambloot, M.J.; Litman, G.W.
Proc. Natl. Acad. Sci. U.S.A. 84, 5868-5872, 1987
A:Title: Extensive families of constant region genes in a phylogenetically primitive vert
A:Reference number: A32716; MUID:87289703; PMID:3475706
A:Accession: C32716
A:Molecule type: DNA
A:Residues: 1-59 <KO2>
A:Cross-references: GB:M17186
A:Genetics:
A:Introns: 100/1; 206/1; 309/1; 419/1; 459/3
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa)
chain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into 1a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
F:120-87/Domain: immunoglobulin homology <IMM1>
F:123-190/Domain: immunoglobulin homology <IMM2>
F:228-291/Domain: immunoglobulin homology <IMM3>
F:330-400/Domain: immunoglobulin homology <IMM4>
F:438-458/Domain: transmembrane #status predicted <TM>
F:27-85,130-188,235-289,337-398/Disulfide bonds: #status predicted
F:164,200,245,374,411,437/Binding site: carbohydrate (Asn) (covalent) #status pr
Query Match 11.2%; Score 301.5; DB 1; Length 461;
Best Local Similarity 26.7%; Pred. No. 1e-10;
Matches 121; Conservative 67; Mismatches 161; Indels 105; Gaps 23;
QY 53 WKNSQIKILNQGSGFLTKGPKSLNDRASRRSLMDQGNFPLIIKNLIKEDSDTYICEV- 111
DB 41 WKDNPEPTTG-----LKTYPSPVLNKKGTQTS-----SGLTTESVGSSKIT-CEVR 88
QY 112 -----EDQKEVQLVFG--LTANSDTHTLLOGSITLT-- 142
DB 89 RGEVWIKELPDCGKDKVHPVTIILTQSSSEITSRFAIVCSIIDFH---PESITVSWL 145
QY 143 -----LESPPGSPSPVQCRSPRGKY--IQGGKILSVSGLQDSGTCTCTVLONOKVEF 195
DB 146 KDQGHMEGFTVSPF--C-----GVNGTSATSRILTPARKEMFTKNVYICQV--SHQGTQ 197
QY 196 KIDI---VPCPAPRPSKCDKTHTCPELLGGPSVFLFPKPKDITMTSRPEVTCVVVD 251
DB 198 SRNITGSGVPC-----SCN-----DPVIKLTPPSIEQVL--LEATYTLTCV--V 237
QY 252 SHEDEPKFENNYVDGVEYHNAKTKPREQYNSTRVSVLTVLHODMLNGEYKCKYSNK 311
DB 238 SNAPYGVNVSW---TQEQKSLKSEIAVQPGEDADSVISTVNIQTQAMLSGAEFYCVVNHQ 294
QY 312 ALPAPIETKISKAKGP--REPOV--YTLPSRDELTKNOVSLCTLVKGFYPSDIAVEMESN 369
DB 295 DLPTPLRASIKKEVKDIREPSVSLTSLSPADVSAQRSLTCLVKGFFPEITVKMTVN 354
QY 370 GQPEN--NYKTTTPVLDSD--GSFPLSKLTVYDKSRMOQGNVFGSVNHEALNHYTOKSL 426
DB 355 DKSVNPKYKNTVEVAENDNSYFIYSLSLIAEEMWAGASVSCVVGHEAL----- 405
QY 427 SLSPGLQDDETCALRQDQDELGLGTLTTPPPRASAL 460
DB 406 ---PLKTIINRTVKNSSDS--DHIMEDNEESAI 435

```

RESULT 85

```

A:Accession: A46507
A:Molecule type: chicken
A:Species: Gallus gallus (chicken)
C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
A:Accession: A46507
R:Mansikka, A.
J. Immunol. 149, 855-861, 1992
A:Title: Chicken Iga H chains. Implications concerning the evolution of H chain genes.

```

```

A:Reference number: A46507; MUID:92340889; PMID:1634774
A:Accession: A46507
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-585 <MAN>
A:Cross-references: GB:S40610; NID:g251907; PID:g251908
A>Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:109906, NCBI:P:109907)
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:34-117/Domain: immunoglobulin homology <IMM>

```

```

Query Match 11.2%; Score 301.5; DB 2; Length 585;
Best Local Similarity 26.8%; Pred. No. 1.4e-10;
Matches 114; Conservative 60; Mismatches 184; Indels 67; Gaps 22;

```

```

QY 19 LPATQGNKVTLLGKGDVETLTCAQSKKSIQFMKNSNQIKILNQGSGFLTKGPKSLND 78
DB 197 LPVATGTGYSL-----TALTVPREQ-----LOGNE--FYCRAQHAATG 234
QY 79 RADSRSLMDQGNFPLIIKNLIKEDSDTYICEVEDQKEVOL--LVFGL--TANSDTHTLQ 135
DB 235 -ADVKEITGGDGVCPITTSKVTLLSDPTQ---EDFERVVLVGLVSGLPAGAAIQWLQ 289
QY 136 -GQSLTTLTSSPSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTVLONOKVE 194
DB 290 DNEBMT-----PAVESDESGSDCTESGVTQWSRVNTTRKSMEGGAQFGCVTHGALK-- 342
QY 195 FKIDIVCPAPRPSKCDKTHTCPELLGGPSVFLFPKPKDITMTSRPEVTCVVVDVSH 254
DB 343 -----EVTYITVSTDCAT--POL---QVSLPPLTEE--LVSHNATVTCVSNMAAA 389
QY 255 DPEYKFMVYVD--GVEYHNAKTKPREQYNSTRVSVLTVLHODMLNGEYKCKYSNK 311
DB 390 D-GVSVWSMSSSGGLV-----SQTEDRQADGRVTASPLRVCAEENNGSTPFGSVREE 444
QY 312 ALPAPIETKISKAKGP--REPOVYTLPSRDELTKNOVSLCTLVKGFYPSDIAVEMESN 369
DB 445 GV-VVAESIRKETDTLHAHSYVFPFPAEELSLQETATLTCAASSFLPSILLTWTQ 503
QY 370 GQP--ENNYTTTPVLDSDSFP--LYSKLTVYDKSRMOQGNVFGSVNHEALNHYTOKSL 426
DB 504 NQPLSPQNYLIFGP--EKDDGFYSLYSKLTVSVEDMQRGDVFGCVVGHGDIPLNFIHKS 561
QY 427 SLSPG 431
DB 562 DKXAG 566

```

RESULT 86

```

A:Accession: S60266
A:Molecule type: novel antigen receptor precursor - nurse shark
A:Species: Ginglymostoma cirratum (nurse shark)
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 21-Jul-2000
A:Accession: S60266
R:Greenberg, A.S.; Avila, D.; Hughes, M.; Hughes, A.; McKinney, E.C.; Flajnik, M.F.
Nature 374, 168-173, 1995
A:Title: A new antigen receptor gene family that undergoes rearrangement and extensive s
A:Reference number: S60266; MUID:55183140; PMID:7877689
A:Accession: S60266
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-684 <GRE>
A:Cross-references: EMBL:U18701; NID:g699442; PID:AB48195.1; PID:g699443

```

```

Query Match 11.2%; Score 301.5; DB 2; Length 684;
Best Local Similarity 25.1%; Pred. No. 1.7e-10;
Matches 107; Conservative 66; Mismatches 161; Indels 92; Gaps 16;

```

```

QY 37 VELTCTASQKSIQFMKNSNQIKILNQG--SFLTKGPKSLNDRASRRSLMDQGNPL 94
DB 321 VETTCASQKQSSSTPVVKTTRKARVETKPHNLPLPSPSEIQTSSATLTCLIRGFDP 380
QY 95 IIRNLKIEDSDTYICEVEDQKEVOLVFGLTANSDTHTLQGGSLTTLTSSPPGSSPSVQ 154

```

```

Db      301 KY-----SVSWQKDVDS--VSANTNPPTAL--EQDLTFSTRS-----414
QY      155 CSRPRKNIQGGKTLTSLQLELDGSGTWCT-----VLQNKQKVEFKIDIVPC 202
Db      415 -----LNLTAIVEMKSGAKYCTASHSPSGSTVRKVRIRNQ-----KVD----452
QY      203 PAPERKSCDKHTCPCLLGGPSVFLPPPKDPTLMISRTPEYTCVVDVSHEDPE-VKRN 261
Db      453 -----CRQTDI-----SVSLKLP-PREEITWQQTATVCEIV--YDDLNIKVF 493
QY      262 WYDGVENVNAKTKPREEQYNSTRYVSVLTVLHODMLNGEKYCKKVSNNKALPAPIEKTI 321
Db      494 WQVNGVERKKGVETQNPENWSGSKSTIVSKLKVMASEMDSGTEYVCLVEDSELPTVYKASI 553
QY      322 SKAK-GQPREPOVYTLPPSRDEL-TKNQVSLTCLYKGFPSDIAVEMESNGO-PENNYKT 378
Db      554 RKANVSQMHPPKCYLHPSTDEIDTENSATLMCLATNFPAPRIYVGMAMNDTLLDSGYRT 613
QY      379 TPVLDSDGSFPLYSKLTVDKSRMOQGNVSSGVMEALH-----NHYTOKSLSLSPG 431
Db      614 QVDSEKSGSSFTVDRLLTLTAEMNSDTTYSCLVGHPSLNRDLINSTKNGKSSVNVVS 673
QY      432 LQLDERT 437
Db      674 VVLSDT 679

```

RESULT 87

```

147175
Ig alpha chain C region - pig (fragment)
C.Species: Sus scrofa domestica (domestic pig)
C.Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C.Accession: I47175
R:Brown, W.R.; Butler, J.E.
Mol. Immunol. 31, 633-642, 1994
A.Title: Characterization of a C alpha gene of swine.
A.Reference number: I47175; MUID:94254897; PMID:7545929
A.Accession: I47175
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-342 <BRO>
A.Cross-references: EMBL:U12594; NID:9555826; PIDN:AAA65943.1; PID:9555827
C.Genetics:
A.Gene: IGACalpha
C.Superfamily: Immunoglobulin C region; immunoglobulin homology
F:232-304/Domain: immunoglobulin homology <IMM>

```

```

Query Match      11.1%; Score 300; DB 2; Length 342;
Best Local Similarity 27.1%; Pred. No. 9e-11;
Matches 96; Conservative 44; Mismatches 138; Indels 76; Gaps 13;

QY      139 LTLTLESPPG-----SSPSVQCRSPRGKNI-----QGGKTLVSQLEL 176
Db      11 LTLGSEPPAGVYVIAVLVDFPSEPLVTWSPSREGVYVRNPPAOGGLVTMSQULT 70
QY      177 QDSGWTCTVLQNKQKVEFKIDIVPCPAPERKSCDKTH-----TCPE 218
Db      71 -----PVEQC PADQILKCVQHLSKSSQSVNVPCKVLPSPDCPQ 109
QY      219 LLGGPSVFLPPPKDPTLMISRTPEYTCVVDVSHEDPEVKRNWYVDGVEVNAKTKPRE 278
Db      110 CC-KKSLSLQPPALAD-LTLGNNASTCTLSGLKSE-GVSTFWPSSGK-DAVQASPR 165
QY      279 EQYNSTRYVSVLTVLHODMLNGEKYCKKVSNNKALPAPIEKTIKAKGPREPOVYTLPP 338
Db      166 DSC-GCVSYVSLTPGADPMNGEFTFSCAASSELKSLATATITPKVTFRPQVHLHP 224
QY      339 SDELTKNQ-VSLTCLYKGFPSDIAVEMESNGO-PENNY--KTPPVLDSDGSFELY 392
Db      225 PEEELALNELVTLTCLVGRFSPKDVLRVWLQGGQELPRKXVYVWESLPEPGQAITVAVT 284
QY      393 SKLTVDKSRMOQGNVSSGVMEALHNHYTOKSLSLSPG-----LQLDERTCAEQ 442

```

```

Db      285 SVLRVADMDWQGDFTSCVGVGHEALPLAFQTKTIDRLAKPHTHVNVVMAEAE 338

```

RESULT 88

```

509264
Ig alpha chain C region - rabbit (fragment)
C.Species: Oryctolagus cuniculus (domestic rabbit)
C.Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C.Accession: S09264
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A.Title: The Iga heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A.Reference number: S09264; MUID:90076124; PMID:2512120
A.Accession: S09264
A.Status: not compared with conceptual translation
A.Molecule type: DNA
A.Residues: 1-339 <BUR>
C.Superfamily: Immunoglobulin C region; immunoglobulin homology
C.Keywords: immunoglobulin
F:124-190/Domain: immunoglobulin homology <IMM>

```

```

Query Match      11.0%; Score 297; DB 2; Length 339;
Best Local Similarity 29.0%; Pred. No. 1.3e-10;
Matches 96; Conservative 48; Mismatches 127; Indels 60; Gaps 15;

```

```

QY      141 LTLSPSSPSVQCRSPRGKNIQ-----GKTLVSQLELDGSGTWCTV 186
Db      10 LSLPCLSGOPVYV-----GCLIQFPPLGVLNVKRTISGENVTPPQDLPTSGLYTSS 64
QY      187 LQNKQKVF-----KID--IYPCPAPERKSCDKHTCPCL-LGGPSVFLPPPK 231
Db      65 LNLTDSECPCTCVACHVNEVDRLILPC-----DTSSCPPTSCGSESLQRPD 117
QY      232 PKDTLMISRTPEYTCVVDVSHEDP-EYKFNWYDGVENVNAKTKPREQYN--STYRV 287
Db      118 LRD-LILGSDASLTCTLGL--KDPKDAVFTW-----EPTNGNEPVQGSPOKDPCCGYSV 169
QY      288 VSVLTVLHODMLNGEKYCKKVSNNKALP-APIEKTIKAKGPREPOVYTLPPSRDELTKN 346
Db      170 SSVLEGCGETYTAGTEPCTVTHPEIESSSLATIRKQGSTLPQVHILPPSEELALN 229
QY      347 Q-VSLTCLYKGFPSDIAVEMESNG-QPENNY--KTPPVLDSDGSFPLYSKLTVDKS 400
Db      230 ALVTITCLVGRFSPKDVLRVWTKNGVVPKPSFLVWKLPEPBGQEPPTVAVTSLRVSAE 289
QY      401 RMOQGNVSSGVMEALHNHYTOKSLSLSPG 431
Db      290 DMNGDSYSVGVGHEGLAHEFTQKTIIDRQAG 320

```

RESULT 89

```

509269
Ig alpha chain C region - rabbit (fragment)
C.Species: Oryctolagus cuniculus (domestic rabbit)
C.Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C.Accession: S09269
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A.Title: The Iga heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A.Reference number: S09264; MUID:90076124; PMID:2512120
A.Accession: S09269
A.Status: not compared with conceptual translation
A.Molecule type: DNA
A.Residues: 1-357 <BUR>
C.Superfamily: Immunoglobulin C region; immunoglobulin homology
C.Keywords: immunoglobulin
F:142-208/Domain: immunoglobulin homology <IMM>

```

```

Query Match      11.0%; Score 296; DB 2; Length 357;
Best Local Similarity 29.8%; Pred. No. 1.6e-10;
Matches 90; Conservative 40; Mismatches 104; Indels 68; Gaps 14;

```

```

OY 141 LTLSPSSSPSVOCRSRGNKNIQCKTSLVSQLELDDSGTWCTVNLQNKVKEPKIDIV 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 94 VTVSP-----EQPPTP-----GSDTTTC----- 115

OY 201 PCPAPBPKCKTKTTCPELLGSPSVLEFPPEKDTLMTSRPEVTCVVVDVSHEDPEYKF 260
Db 116 PCPCPSP-SC-----GEPSSLORPFLRD-LTLNSNASLITCLTRGKNPEGAN-F 162

OY 261 NMVYDGVENVHNAKTKPREEQYNS----TYRVSVTVTLVHODMLNGKEYCKVSKNALD-A 315
Db 163 TW-----EPTNG-NKPVQGSVQSYPCGCYSVSVLPGCAEFMNNGTEFTCTVTHPEIDGG 216

OY 316 PIEKTIKAKQKPREPQVYTLPPSRDELTKQ-VSLTLVYGFPSDIAYMESNG--QP 372
Db 217 PLTAKISDGTALIPQVHLLPPSEELTALNELVTLTCLVYGFSPKDLVYMTKGVNVP 276

OY 373 ENNY---KTPPVLDSDSFFLYSKLTVDKSRMOQGNFSCSVMEALHNNHYTOKSLIS 429
Db 277 ENSFLVWKPLPEPQGEPTTAVTSLRLVPAEDMNMNESYTCVVGHEGLAEHFTQRTIDRL 336

OY 430 PG 431
Db 337 SG 338

```

Query Match	Similarity	Score	DB 1;	Length	367;
Best Local	85;	Conservative	48;	Mismatches	119;
				Indels	30;
				Gaps	10;
160	GKNIQGGKTLTSSVLELQDSTGWTCTVLQNGKVEFKIDIVCPAPEKSCDKHTC--	216			
87	GPEVVASRISVTSSEMDTGATFECV-----EGEKRNATSKMECGLE	129			
217	PELIGGDSVLEFPKPKDITMIISRTPEVTCVVDVSHEDP-EVKNFNVYDGVENHAKTK	275			
130	PVVGQDDAIRVITPESFVD-IFISKATLTICVSNMNVNADGLEVSM-WKEGGKLETALGK	187			
276	PREQGVNSTRVNVSVLYLHODMLNGKEYKKKVSNNKALPAIEKTISKAK-GQREPOVY	334			
188	RVLQSGNLTVDDVAATVVCASEWDGGGYVCVKVHPDLLFEMEKMKRTKXASNARPSVY	246			
335	TLPPSRDELTKNQ-VSLTCLVAKGVPSDIAVENESNGQ--PENNYKTTPTPLVDS--DGSF	389			
247	VFPPTPTQLANGNQSLSTTCMAQGFNPRLFPKWRKNGEPLPQSGSVTSAPMAENPENESY	306			
390	FLYSKLTVDKSRWQGVNFCGSVMHEALHNHYTOKSLSLSPG	431			

```

Db          307 VAISVVGAEEMGAGNYTCLVGHEALPLQLAKQKSVDRASG   348
|||||      |||||      |||||      |
RESULT 91
A46532
Ig mu chain C region - axo1ocl (fragment)
C1Species: Ambystoma mexicanum (axo1cl)
C1Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C1Accession: A46532
R1Fellah, J.S.; Miles, M.V.; Charlemagne, J.; Schwager, J.
Eur. J. Immunol. 22, 2595-2601, 1992
A1Title: Evolution of vertebrate IgM: complete amino acid sequence of the constant region.
A1Reference number: A46532; MUID:93011455; PMID:1382992
A1Accession: A46532
A1Status: preliminary
A1Molecule type: nucleic acid
A1Residues: 1-454 <FEU>
A1Note: sequence inconsistent with the nucleotide translation
C1Superfamily: immunoglobulin C region; immunoglobulin homology
C1Keywords: immunoglobulin

```

```

Query Match Summary      11.0%; Score 296; DB 2; Length 454;
Best Local Similarity 23.1%; Pred. No. 2, 2e-10;
Matches 111; Conservative 65; Mismatches 190; Indels 114; Gaps 17;

QY      21 AATGQNKV---LGGKDYELTCTAS--QKKSIOFHMKNSQIKI--LGNQSFLLTK 71
Db      1 AAAQAPSVYPLISCASLDPPVIGCLAKGFLPDSVTFEDMTDKNNASFSAGVAKLPSTVTTG 60
QY      72 GPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDPTICEFEDQKEFQQLVFGLTANSPT 131
Db      61 GLYSASTVVKVPSEW-----KARDPYCKV-----T 87
QY      132 HLLOGSITLTLLESPDG--SSPSVQCRSP-----R 159
Db      88 HPSLQQLPTKKVQFVSQIRSKPTVTLHAAPARDIINNATVTCICRGFHPQIPISKMMKN 147
QY      160 GKNIQGGKTLVSQLELDQSGTWCTCTVLONQKKVEKIDIV-PCPAPE-----PKSCD 211
Db      148 GKDVLTSG--ITTEBEVADTAGNFVDVTSLNTIEPMDNMMDTVVSCVVDQTSKFWNTNMS 205
QY      212 KTHTCPELLGGPSVFLPPKPK-DTLMTSRTEVTCVVVDV-----SHEDPEVKF 260
Db      206 KSMLCDQAVGEPVKTATFTVAFTFEDMFESKSNANTVCIYNNMGTIEGPNITWSREDTNEVL 265
QY      261 NMVYDGEVHNAAKTPREEQVNSTYRVVSVLTVLHODWLNKGEYCKCKVSNKALPAPIEKT 320
Db      266 KTEITNPIFHD-----NATLSVMGIATVCAQDMQDANHFKFVCVLHQDLAEQEVLS 315
QY      321 ISKAKG-QPREPQVYTLPPSRDELTKNOVSLTCLVKGYPYSLIAVMEBSNGQ--PENNY 376
Db      316 LQKPNGNRRKRPVSVIYPPPSSEELAKETATVTCILMRGHPCDLFRMLIENSQLOQODY 375
QY      377 KTEPPVLDS-----GSFFLYSKLTVDKSRMQQGNVFSVMEALVHNHYTOKSLSLSPG 431
Db      376 VNTQAEBEVDPTTGKSCFMYSMLKIPAAQWTAAGNTYTCVGHGHALPLOTQKSIDRSFG 435

RESULT 92
S09272
IG alpha chain C region - rabbit (fragment)
CISpecies: Oryctolagus cuniculus (domestic rabbit)
CIDate: 29-Jan-1993 #sequence_revision 29-Jan-1993 #ext_change 16-Jul-1999
CIDAccession: S09272
EMBO J. 8, 4041-4047, 1989
R;Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
A;Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A;Reference number: S09264; MUID:90076124; PMID:2512120
A;Accession: S09272
A;Status: not compared with conceptual translation
A;Molecule type: DNA

```

A:Residues: 1-343 <BUR>
C:Superfamily: Immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:232-305/Domain: immunoglobulin homology <IMM>

Query Match 10.9%; Score 295.5; DB 2; Length 343;
Best Local Similarity 31.7%; Pred. No. 1.7e-10;

Matches 96; Conservative 42; Mismatches 120; Indels 45; Gaps 16;

QY 146 PEGSSSVCCRRPRKGNIGGKTLISVQLELDSGWTCTVLQNKVFEKIDI-VPCP- 203

DB 50 PVPSPSSSLYTT-----SSLSTLDQCPRDGVNTHVHNVDGQ---DLTVPCOD 99

QY 204 ----APEPKCDKTHTCPELGGSPVFLFPKPKDITMISRTPEVYCVVVDVSHDEP-V 258

DB 100 CHCYCP-PTSC-----GEPLSLQRPDIDGLLESKA-SLCTLSGL--KDPBGA 145

QY 259 KENMY-VDSVEVHNATKPREEOYN-STYRVVSVLTVLHODMLNGKEYCKVSNKALP-A 315

DB 146 VFTWEPTNG----NEFVQSVQSVPCGCVSVSLPGCAEPNNAETFTCTVTHPEIEGG 201

QY 316 PLEKTSKAKGQPREQVYTLTPPSRDELTKN-QVSLTCLVKGFPYEDIAVEMESNGQ--- 371

DB 202 SLTATISIRGSLTPQVHLPPTEELALNEQVTLTCLVQGRSPKDVLSVTHNGTLVY 261

QY 372 PENNY---KTPPVLDSDGSPFLYSKLTVDKSRMOQGNVFGSVMEALHNHYTKSL 428

DB 262 FKDSYLVMKPLPEPGODPTTAVTSLIRVSAEDMNGDSYSCVGHGELAEHPTOKTIDR 321

QY 429 SPG 431

DB 322 QAG 324

RESULT 93

B22360 Ig alpha-2 chain C region (allotype A2m(1)) - human

C:Species: Homo sapiens (man)

C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999

C:Accession: B22360

R:Flanagan, J.G.; Lefranc, M.P.; Rabbitts, T.H.

Cell 36, 681-688, 1984

A:Title: Mechanisms of divergence and convergence of the human immunoglobulin alpha-1 an

A:Reference number: A94653; MUID:84130179; PMID:6421489

A:Accession: B22360

A:Molecule type: DNA

A:Residues: 1-340 <FLA>

C:Genetics:

A:Gene: GDB:IGHA2

A:Cross-references: GDB:119333; OMIM:147000

A:Map position: 14q32.33-14q32.33

A:Introns: 1/1 103/1; 210/1

C:Superfamily: Immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:230-302/Domain: immunoglobulin homology <IMM>

Query Match 10.9%; Score 294.5; DB 2; Length 340;
Best Local Similarity 29.2%; Pred. No. 1.9e-10;

Matches 103; Conservative 50; Mismatches 143; Indels 57; Gaps 18;

QY 122 VEGLTANSDTH-----LGG---QSLLTLLESPPSSPSVQGRS-PRGNIOGKGT 168

DB 8 VPPPLSDSTPQDGNVYVACLVOGFPPQEPUSVTWSE---SGQVNTAKNPPSDASGDLY 64

QY 169 LSVSQUEL-----QDSGTWCTVLQNKVFEKIDI-VPCAPPEPKSCDKTHTCPELGG 222

DB 65 TTSSTQLTLATQCPDCKSVTCHVHTNPSQ---DVTVCVPVPPPPC-----C-----H 111

QY 223 PSVFLPEPKPKDTLMSRTPEVTCVVVDVSHDEPVEKFNMYVDGVEVHNATKPREEOYN 282

DB 112 PRLSLRPALED-LTLAGSEANLCTLTGJ-RDASGATFTWPSGK--SAVQGPPEPRLD 167

QY 283 STYRVVSVLTVLHODMLNGKEYCKVSNKALPAPLEKTSKAKGQPREQVYTLTPSRDE 342

DB 168 GCYSVSVLPGCAQPMNGETFTCAAPBELKPLTANITKS-GNTPREVALPPEPSE 226

QY 343 LTKNQ-VSLTCLVKGFPYEDIAVEMESNGQ--PENNYKTPVLD-SG--SFLYSKLT 396

DB 227 LALNELVLTLCIARFSPDVLVRLVQSQELPRKRYLTWASROBPQGTTFVATSLIR 286

QY 397 VDKSMOQGNVSGSVMEALHNHYTKSLSPG-----LQLEDC 438

DB 287 VAEDMKKGDITSQVGHGELPLAFTOKTIDRLAGKPHVNVSVMAEVDGIC 339

RESULT 94

HRKKS

Ig mu chain C region, secreted (clone 3050) - horn shark

C:Species: Heterodontus francisci (horn shark)

C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999

C:Accession: S01853

R:Kokubun, F.; Hinde, K.; Litman, R.; Shambloot, M.J.; Litman, G.W.

EMBO J. 7, 1979-1986, 1988

A:Title: Complete structure and organization of immunoglobulin heavy chain constant regi

A:Reference number: S00980; MUID:88328985; PMID:3138109

A:Accession: S01853

A:Molecule type: DNA

A:Residues: 1-438 <KOK>

A:Cross-references: EMBL:X07781

A>Note: the sequence was determined from the germline gene

C:Genetics:

A:Introns: 100/1; 206/1; 309/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kai

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1;

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob

F:1-438/Domain: C region <CRE>

F:20-87/Domain: immunoglobulin homology <IM1>

F:123-190/Domain: immunoglobulin homology <IM2>

F:228-291/Domain: immunoglobulin homology <IM3>

F:330-400/Domain: immunoglobulin homology <IM4>

F:164,200,245,275,374,411,415,425/Binding site: carbohydrate (Asn) (covalent) #status p

Query Match 10.9%; Score 294; DB 1; Length 438;
Best Local Similarity 27.1%; Pred. No. 2.8e-10;

Matches 115; Conservative 65; Mismatches 153; Indels 92; Gaps 21;

QY 53 WNSNQIKLNGQSFLLTKGPEKLNDRADSRSLMDQGNFLLIKLKEEDDTTCY- 111

DB 41 WKDNEPITTG---LKTYPVLNKKGTYYTQ---SQLTISEVSGSKTY-CEVR 88

QY 112 -----EDKKEVQLLVFG--LTRANSPTHLOGOSLTLT-- 142

DB 89 RGEVWIKELIPCKKDKVNPVYILTQSSSEETSSRFATVLCSTIDFH--PESTIVSWL 145

QY 143 -----LESPGSSPSPVQGRS-PRGN-IOGKTLISVQLELDSGWTCTVLQNKVFE 195

DB 146 KDGQMGSEGFVSPF--C-----GVNGTSPATSRLLVPAREWNTNKVYTCQV--SHQGVYQ 197

QY 196 KIDI---VPCAPPEPKSCDKTHTCPELGGSPVFLFPKPKDTLMSRTPEVTCVVVDV 251

DB 198 SHNITGQVPC-----SCN-----DEVIKILPPSIEQVL-LKATVLTTCV--V 237

QY 252 SHEDDEVKFNMYVDGVEVHNATKPREEOYNSTYRVVSVLTVLHODMLNGKEYCKVSNK 311

DB 238 SNAAPGVNVSW---TOEQKSLSEIAVQGEDADSVISTVNSTQAMLSGAELFYCVVNHQ 294

QY 312 ALPAPIEKTSKAKGP-REPOV-YTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESN 369

DB 295 DLPFLRASIHKEEVYDAREPSVSLISPAEDVSAQRFLSLTCLVKGFPREIPIKMTWN 354

QY 370 GQPEN--NYKTPPVLDSD-GSFLYSKLTVDKSMOQGNVSGSVMEALHNHYTKSL 426

DB 355 DKSVPNGVNTKTEVNAENDNSYFTYSLISIAEEMASGASYSQVGHGELPLKTIINRTV 414

QY 427 SLSPG 431

Db 415 NKSSG 419

RESULT 95

IG alpha chain C region - rabbit (fragment)
C|Species: Oryctolagus cuniculus (domestic rabbit)
C|Date: 28-Aug-1995 #sequence_revision 28-Aug-1995 #text_change 16-Jul-1999
C|Accession: A02174
R|Knight, K.L.; Martens, C.D.; Stoklosa, C.M.; Schneiderman, R.D.
Nucleic Acids Res. 12, 1657-1670, 1984
A|Title: Genes encoding alpha-heavy chains of rabbit IgA: characterization of cDNA encod
A|Reference number: A02174; MUID:84144059; PMID:6322114
A|Accession: A02174
A|Molecule type: mRNA
A|Residues: 1-299 <GB>
A|Cross-references: KIM>X00353; NID:G1575; PIDN:CA25100.1; PID:G1576
C|Comment: This immunoglobulin belongs to the IgA-g subclass. It was isolated from a rab
C|Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa)
C|chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into d
C|Superfamily: immunoglobulin C region; immunoglobulin homology
C|Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin; plasma
E|86-153|Domain: immunoglobulin homology <IM1>
E|189-261|Domain: immunoglobulin homology
E|38,286|Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match	10.8%	Score 293	DB 1	Length 299
-------------	-------	-----------	------	------------

QY 200 VPCDAEPEKSCDHTKTCP-----ELLGG--PEVFLPEPRKPDTLMIISTPEVTCVVVDV 252

Db 41 LPPFPFDCCBPANCCCTCPSSSSRNLLISGCGPRLSGQRDLGD--LLGRDASLTITLSGLK 99

QY 253 HEDPEVKKFMWYDGVENVNAKTKPRBEQYNSTYRVVSLVYLHODMLNGKCYKCKYSNKA 312

Db 100 NPEDAA-FTW--EPTNGNEPVOOGRADRDLSGCYSVSVLPSSAETWKARTEFTCTVTHPE 156

QY 313 LPA-PIEKTIKSAKQGRPREPOVYTLPRSHDELTKR-QVSLCYLKGFYPSDIANEMWSNG 370

Db 157 IDSGSLTATISKCVLP--PQVHLPRPEEELALBEQVTLTCLVARGSPKQVLVSMNHG 214

QY 371 Q--PENNY---KTPPEVLDDSGSPFLYSKLTVDKSRMQQGVNFCSSVNHGALHNHYOKS 425

Db 215 QEVPEDSPLVWKSMPRESSQDKATYATLSLRLRPADDMVQGDYLSGVMGHBEALAHFQKT 274

QY 426 LSLSPG 431

Db 275 IDRLAG 280

RESULT 96

IG alpha chain C region - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S09276
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A:Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A:Reference number: S09284; MUID:90076124; PMID:2512120
A:Accession: S09276
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-338 <BUR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
C:128-300/Domain: immunoglobulin homology <IMM>

Query Match	10.8%	Score 292;	DB 2;	Length 338;
Best Local Similarity	29.2%;	Pred. No. 2.7e-10;		
Matches	92;	Conservative	50;	Mismatches 133;
			Indels	40;
			Gaps	15

```

OY      146  PESSSVQCRSPRGKNIQGGKTTLSVQLDELQDSEWTWC-----TVLQNOCK-----V 133
Db      16  PPG--PLVSWVTNGENV----SAYNFPFPAQSGTSGPYTACSELLPLVTQCLGQGSAAACHV 70
OY      194  EFKIDI---VPCAPPEPKSCDKTHNCP-----ELLGG--PSVFLFPPEKXDTLMIISRPE 243
Db      71  EYNSVINEBLSPYVFPFPPCCPANSCTCTPSSSRNLISGQFSLSLQRLDLD--LLGGDAS 129
OY      244  VTCVVVDVSHEDPEYKFNMYVDGVEVHNNAKTRPREQYNSTYRVVSVLTVLHQDI1MLGKE 303
Db      130  LTCTLSGLKNRPEDAV--FTW--EPTGNENPVGQQRQDLDSGCVSVSVLPSAEFTWKARTE 186
OY      304  YKKVSNKKLPLA-PIEKITSKAKGQPREVOYTLTPRSDELTKA-QVSLTGLVKGF1PSD 361
Db      187  FCTVTTHPRIDSGSLTATISRGVLP--POYHLLPRPESEELALNEQVTLTCLVGFSG1PKD 244
OY      362  IATWEMSNQ--PENNY---KTPPVLDSDGSFLYSLTLTDKSRMQGNAV1FGSVNHEA 416
Db      245  VLVSMWHQOEVEEDSFLVWKSMPRESSODKATVATISLPLRVADMDNQD1YSCMVHEG 304
OY      417  LHHNYTKOSSLSPG 431
Db      305  LAEHFTQRTIDRLAG 319

```

RESULT 97

Ig alpha chain C region - rabbit (fragment)
C|Species: Oryctolagus cuniculus (domestic rabbit)
C|Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C|Accession: S09286
R|Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A|Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A|Reference number: S09284; MUID:90076124; PMID:2512120
A|Accession: S09286
A|Status: not compared with conceptual translation
A|Molecule type: DNA
A|Residues: 1-352 <BUR>
C|Superfamily: immunoglobulin C region; immunoglobulin homology
C|Keywords: immunoglobulin
F|241-314/Domain: immunoglobulin homology <IMN>

Query Match	10.8%;	Score 292;	DB 2;	Length 352;
Best Local Similarity	31.2%;	Pred. No. 2.8e-10;		
Matches 95;	Conservative 41;	Mismatches 118;	Indels 50;	Gaps 16

```

Oy      139  LTLTLSSPGSSPSVQCSPRKGNIGGKTLTSSVLSLELDSDSGTTCVTLNQNKVVERKID 198
Db      69  LSLTPHQCP-EDHNVAVCVEH--NPNDEQONLTIVLYPECKDSS-----D 109
Oy      199  IVCPAPAEPEKSCDKHTTCPELLGSGSVFLPPPKEDTLMTSRTPBVTGVVVDVSHEDPE 257
Db      110  PTCPCPP-PITC-----GEPSLSLQRPDIDG-LLESNASTLTLSGL-KDREG 155
Oy      258  VKENMY-VDGEVHNNAKTKPREEOYN-STYEVSVSLTVLHODWLNKREYCKCKSNKALP 314
Db      156  AVFTWMPITNGNEFVQOST----QSVPCGCSVSVSLPECAEPMAAGIEFTCTVTHREIEG 211
Oy      315  APIEKTISKAKOPREPOVYTLPRSEDLTYNQ-VSLTCLVKGFPSPSDIAVEWESNCO- 371
Db      212  GSLTATISRGLTP--PQVHLHPRTEELALNLTLCVLRGSPSPDVLVSWTHNGTIV 269
Oy      372  -PENNY---KTPPVLDSDGSEFLYSKTLVDKSNMOGQSNFSGSVMEALHNHTYQSL 427
Db      270  VPKDSFLWKMPLEPGOEFTTYAVTSLRVAEDMNOQDSYSYCVGHGEGALHEHTORTID 329
Oy      428  LSPG 431
Db      330  RLNG 333

```


RESULT 98

A34891

Ig heavy chain precursor V region - ladyfish

C:Species: Elops saurus (ladyfish)

C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #ext_change 21-Jul-2000

C:Accession: A34891

R:Amemiya, C.T.; Litman, G.W.

Proc. Natl. Acad. Sci. U.S.A. 87, 811-815, 1990

A:Title: Complete nucleotide sequence of an immunoglobulin heavy-chain gene and analysis

A:Reference number: A34891; PMID:9018916; PMID:2105490

A:Accession: A34891

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-568 <AME>

A:Cross-references: GB:M26182; NID:g213134; PIDN:AAA49238.1; PID:g213135

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: heterotrimer; immunoglobulin

F:33-116/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 22.3%; Pred. No. 6.5e-10;

Matches 92; Conservative 78; Mismatches 152; Indels 90; Gaps 15;

QY 65 QGSFLTKGPKSLNDRADSRSLMDQGNFPLI-----KNLKIED 103

DB 207 EGDFKSKAEKCKTELGGKKT-----PVIKPKPKPKPPQPVLSIMTSGEELTLNK 259

QY 104 SDTYIEVEDQKEVQLVFGLTANSDTNLLQGSGLTLESPPSSPSVQCRSPKKN 163

DB 260 TATFALCLAND-----FYPKGHSFKMLRDGEVLT-----DGIALTLTECKKKDKSF 304

QY 164 QGSKTLVSQLELD-SGTMCTVLNOKKVE--FKIDVPCAPAPKSDCKHTPELL 220

DB 305 TASSFLQASBSQKRLDGLFTCCFQEGETTEVTAKSAC-SPNAQIDAK----- 355

QY 221 GGSFVLFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFMYVDGVE--VHNATKPKPE 278

DB 356 -----ISPPTPEE-LFLQGTTRLTC-----KITGVDDGVNVTMEVSGEAVV 396

QY 279 EQYNSTYRVVSVLTVLHODWLNGEKYCKVSKNAPLEKITSK-AKQGPPEPVYTL 337

DB 397 GGPDE-QKMSKLLIYEEKNRTEYTCVKEHSDLSPLTSYRRCGSGKMGSPFVFLA 455

QY 338 PSRDELTKQVSLTCLVKGFYPSDIAVEMESNGEPENNTKTPPVLDSSGFFLYSKLT 397

DB 456 PA-EQNNLSVLTCLCAKQDTPFQVLIWLVDQPEVTPTEVTKTEGTSVFSQLTI 514

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

DB 515 PASDWGSGVYVSCAVNHETMESVNTVTRTDSVSKKPTTVSLDLNVPQTC 566

QY 398 DKSRMOQGVNFCSCVNHAEAL-----HNHYTKSLSLSPGLQDDETC 438

QY 134 LOGOSLTU-TLESPPGS-----SPSVOCR---SPRKNIOGKTLISVS 172

DB 43 VSGSEVTPPPVSPSSSLYTTYSLLRLPAEQCPRENSVACRBNHKKGDVVPSPACN 102

QY 173 QLELDGSGTMTCTVLONOKKVEFKIDVPCAPAPKSCDKHTCEBLLGSPVLPKPK 232

DB 103 ESTIEPTPTTC-----PCPEPSP-SC-----GKPLSLQRPDL 135

QY 233 KDTLMISRTPEVTCVVVDVSHEDPEVKFMYVDGVEVHNATKPKPE--QNSTYRVVS 289

DB 136 GD-LILNSNASLTCTLRGLNPEGAV-FTW-----EPFPGKSPVQGSPLDHCVCYSVS 188

QY 290 VLTVLHODWLNGEKYCKVSKNAPLEKITSKAKGPRPEPVYTLPPSRDELTKNQ- 347

DB 189 VLPGCATVNNAGTEFTCTVTHPEIGDSLTGTSKDSGLPPOVHLPPSEBELALML 248

QY 348 VSLTCLVKGFYPSDIAVEMESNGQ--PENNYKTPPVLD--SDGSFLYSKLTVDKSR 401

DB 249 VTLTCLVRFSPKDVLSVTHNGTVPVPRKDSVLWKPLREPODPTTVAITSLRVPAD 308

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

QY 402 MOQGVNFCSCVNHAEALHNHYTKSLSLSPG 431

DB 309 KNOGDSYSCVVGHEGLAEHFTOKTIDRLAG 338

Db 287 VAEDMKKGDTPSCWVGHREALPLAFTQKTIDRLAGKPTHVNSVYMAEVDGTC 339

Search completed: August 3, 2004, 13:15:19
Job time : 16.7668 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 13:01:34 ; Search time 19.0695 Seconds
(without alignments)
1754.300 Million cell updates/sec

Title: SEQ07
Perfect score: 3414
Sequence: 1 MNRGVPRRLTLVLALLP.....DETCAEAQDELGLMTTDP 648

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 125 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	3189	93.4	630	4	US-08-472-888A-6
2	2116	62.0	530	3	US-08-477-460B-4
3	2116	62.0	530	3	US-08-379-516-4
4	2116	62.0	530	3	US-09-329-916-4
5	2116	62.0	530	3	US-08-485-372A-4
6	2116	62.0	530	4	US-09-409-006A-4
7	2116	62.0	530	4	US-08-484-681-4
8	2116	62.0	530	5	PCT-US93-07422-4
9	2059	60.3	432	3	US-08-477-460B-2
10	2059	60.3	432	3	US-08-379-516-2
11	2059	60.3	432	3	US-09-329-916-2
12	2059	60.3	432	4	US-08-485-372A-2
13	2059	60.3	432	4	US-09-409-006A-2
14	2059	60.3	432	4	US-08-484-681-2
15	2059	60.3	432	5	PCT-US93-07422-2
16	2036.5	59.7	532	2	US-08-417-485-6
17	2036.5	59.7	532	2	US-08-284-391B-6
18	2036.5	59.7	532	3	US-09-218-950-6
19	2036.5	59.7	532	5	PCT-US92-01785-6
20	2036.5	59.7	532	5	PCT-US95-00454-6
21	2036.5	59.7	575	2	US-08-417-495-4
22	2036.5	59.7	575	2	US-08-284-391B-4
23	2036.5	59.7	575	3	US-09-218-950-4
24	2036.5	59.7	575	5	PCT-US92-01785-4
25	2036.5	59.7	575	5	PCT-US95-00454-4
26	2032.5	59.5	462	2	US-08-417-495-5
27	2032.5	59.5	462	2	US-08-284-391B-5

28	2032.5	59.5	462	3	US-09-218-950-5	Sequence 5, Appli
29	2032.5	59.5	462	5	PCT-US92-01785-5	Sequence 5, Appli
30	2032.5	59.5	462	5	PCT-US95-00454-5	Sequence 5, Appli
31	2029	59.4	398	2	US-08-284-391B-29	Sequence 29, Appli
32	2029	59.4	398	2	US-09-218-950-29	Sequence 29, Appli
33	2029	59.4	458	3	US-08-466-368-4	Sequence 4, Appli
34	2021	59.2	457	4	US-08-328-500-9	Sequence 9, Appli
35	2017	59.1	402	1	US-08-236-311-1	Sequence 1, Appli
36	2017	59.1	402	3	US-08-457-918-1	Sequence 1, Appli
37	2015	59.0	458	3	US-09-039-555B-15	Sequence 15, Appli
38	2007	58.8	458	3	US-09-517-605-3	Sequence 3, Appli
39	2001	58.6	394	3	US-08-466-368-2	Sequence 2, Appli
40	2001	58.6	394	4	US-08-328-500-2	Sequence 2, Appli
41	1989	58.3	394	6	5223394-7	Patent No. 5223394
42	1951	57.1	394	6	5223394-9	Patent No. 5223394
43	1904	55.8	434	1	US-08-226-311-4	Sequence 4, Appli
44	1904	55.8	434	3	US-08-457-918-4	Sequence 4, Appli
45	1892	55.4	433	2	US-08-867-149-1	Sequence 1, Appli
46	1892	55.4	433	2	US-08-808-374-1	Sequence 1, Appli
47	1887	55.3	433	3	US-09-100-409A-1	Sequence 1, Appli
48	1695	49.6	433	6	5171838-13	Patent No. 5171838
49	1585.5	46.4	410	3	US-08-630-172-17	Sequence 17, Appli
50	1585.5	46.4	410	3	US-09-375-419-17	Sequence 17, Appli
51	1368	40.1	318	6	5223394-11	Patent No. 5223394
52	1363	39.9	295	6	5223394-9	Patent No. 5223394
53	1338.5	39.2	254	2	US-08-284-391B-33	Sequence 33, Appli
54	1338.5	39.2	254	3	US-09-218-950-33	Sequence 33, Appli
55	1275	37.3	592	4	US-09-313-942-8	Sequence 8, Appli
56	1243	36.4	622	4	US-09-499-846-2	Sequence 2, Appli
57	1237	36.2	704	4	US-09-590-656-2	Sequence 2, Appli
58	1237	36.2	704	4	US-09-733-764-2	Sequence 2, Appli
59	1235.5	36.2	497	4	US-09-499-846-6	Sequence 6, Appli
60	1235.5	36.2	525	4	US-09-499-846-4	Sequence 4, Appli
61	1232.5	36.1	424	5	PCT-US95-03866-12	Sequence 12, Appli
62	1232.5	36.1	424	5	PCT-US95-03866-14	Sequence 14, Appli
63	1232.5	36.1	459	1	US-08-157-101A-7	Sequence 7, Appli
64	1230.5	36.0	446	3	US-08-397-411-7	Sequence 7, Appli
65	1225.5	35.9	488	3	US-08-776-511-2	Sequence 2, Appli
66	1225	35.8	488	4	US-09-499-846-12	Sequence 12, Appli
67	1223.5	35.8	680	3	US-08-227-996C-15	Sequence 15, Appli
68	1222.5	35.8	454	2	US-07-934-733C-22	Sequence 22, Appli
69	1222.5	35.8	454	3	US-08-437-642B-22	Sequence 22, Appli
70	1222.5	35.8	454	4	US-08-146-206C-22	Sequence 22, Appli
71	1222.5	35.8	454	4	US-09-705-686-22	Sequence 22, Appli
72	1222.5	35.8	454	5	PCT-US93-07832-22	Sequence 22, Appli
73	1221.5	35.8	475	4	US-09-740-002-25	Sequence 25, Appli
74	1221.5	35.8	473	3	US-09-049-672A-4	Sequence 4, Appli
75	1220.5	35.7	387	1	US-08-470-299-4	Sequence 4, Appli
76	1220.5	35.7	453	4	US-09-532-856-6	Sequence 6, Appli
77	1220.5	35.7	453	4	US-09-524-100C-6	Sequence 6, Appli
78	1220.5	35.7	497	4	US-09-499-846-10	Sequence 10, Appli
79	1220.5	35.7	525	4	US-09-499-846-8	Sequence 8, Appli
80	1220	35.7	388	3	US-09-131-247-16	Sequence 16, Appli
81	1219.5	35.7	453	4	US-09-301-593-18	Sequence 18, Appli
82	1219.5	35.7	475	4	US-09-740-002-27	Sequence 27, Appli
83	1219.5	35.7	476	3	US-08-487-550-12	Sequence 12, Appli
84	1219.5	35.7	476	4	US-09-526-098-12	Sequence 12, Appli
85	1216	35.6	442	4	US-08-472-888A-7	Sequence 7, Appli
86	1216	35.6	442	5	PCT-US96-10043-9	Sequence 9, Appli
87	1215.5	35.6	437	3	PCT-US96-10043-11	Sequence 11, Appli
88	1215	35.6	476	2	US-08-378-939-11	Sequence 10, Appli
89	1213.5	35.5	446	4	US-08-157-452B-12	Sequence 12, Appli
90	1213	35.5	449	1	US-08-458-516-13	Sequence 13, Appli
91	1212.5	35.5	468	4	US-09-465-378B-67	Sequence 67, Appli
92	1212.5	35.5	711	4	US-09-485-378B-90	Sequence 90, Appli
93	1212.5	35.5	859	4	US-09-313-942-7	Sequence 7, Appli
94	1212.5	35.5	951	4	US-09-313-942-9	Sequence 9, Appli
95	1212	35.5	451	4	US-09-247-352-3	Sequence 3, Appli
96	1212	35.5	451	4	US-09-466-635-3	Sequence 3, Appli
97	1212	35.5	472	4	US-09-301-593-30	Sequence 30, Appli
98	1211.5	35.5	232	2	US-08-595-043A-50	Sequence 50, Appli
99	1211.5	35.5	331	3	US-09-178-669-2	Sequence 2, Appli
100	1211.5	35.5	331	4	US-09-761-413-2	Sequence 2, Appli

```

101 1211.5 35.5 360 4 US-09-180-100-11 Sequence 11, Appl
102 1211.5 35.5 371 1 US-08-236-111-7 Sequence 7, Appl
103 1211.5 35.5 371 3 US-08-457-918-7 Sequence 7, Appl
104 1211.5 35.5 376 4 US-09-180-100-22 Sequence 22, Appl
105 1211.5 35.5 366 2 US-08-784-512-3 Sequence 3, Appl
106 1211.5 35.5 366 3 US-09-176-228-3 Sequence 3, Appl
107 1211.5 35.5 422 3 US-09-027-449-71 Sequence 71, Appl
108 1211.5 35.5 422 3 US-09-026-985-71 Sequence 71, Appl
109 1211.5 35.5 422 4 US-09-121-952A-71 Sequence 71, Appl
110 1211.5 35.5 426 4 US-09-234-340A-71 Sequence 71, Appl
111 1211.5 35.5 476 3 US-08-487-550-4 Sequence 4, Appl
112 1211.5 35.5 476 4 US-09-526-098-4 Sequence 4, Appl
113 1211.5 35.5 478 3 US-08-487-550-8 Sequence 8, Appl
114 1211.5 35.5 478 4 US-09-526-098-8 Sequence 8, Appl
115 1211.5 35.5 547 4 US-09-746-359A-54 Sequence 54, Appl
116 1211.5 35.5 571 4 US-09-746-359A-53 Sequence 53, Appl
117 1211.5 35.5 450 4 US-09-532-856-5 Sequence 5, Appl
118 1211.5 35.5 450 4 US-09-524-100C-5 Sequence 5, Appl
119 1211.5 35.5 472 4 US-09-301-593-43 Sequence 43, Appl
120 1210 35.4 451 2 US-08-887-352B-18 Sequence 18, Appl
121 1210 35.4 451 3 US-09-109-207C-18 Sequence 18, Appl
122 1210 35.4 451 3 US-09-282-505-2 Sequence 2, Appl
123 1210 35.4 451 3 US-09-054-255-2 Sequence 2, Appl
124 1210 35.4 451 3 US-09-296-005-18 Sequence 18, Appl
125 1210 35.4 451 4 US-09-282-846-2 Sequence 2, Appl

```

ALIGNMENTS

RESULT 1

```

US-08-472-888A-6
; Sequence 6, Application US/08472888A
; Patent No. 6613746

```

```

; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Walz, Gerd
; TITLE OF INVENTION: AGP-ANTIBODY FUSION PROTEINS
; TITLE OF INVENTION: AND RELATED MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,888A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/618,314
; FILING DATE: 23-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/258001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein
US-08-472-888A-6

```

```

Query Match 93.4%; Score 3189; DB 4; Length 630;
Best Local Similarity 97.8%; Pred. No. 5,1e-245;
Matches 617; Conservative 0; Mismatches 6; Indels 8; Gaps 4;

```

```

QY 1 NMRGVPFRLILVLTALIPATQGNKYVLGKGGPTVELTCTASQKSIQFMKNSNQK 60
DB 1 NMRGVPFRLILVLTALIPATQGNKYVLGKGGPTVELTCTASQKSIQFMKNSNQK 60
QY 61 IINGGGSFLLTKPSKLNDRADSRSLMDQGNPFLIINKLIRDSPTYICEVEDQKEVOL 120
DB 61 IINGGGSFLLTKPSKLNDRADSRSLMDQGNPFLIINKLIRDSPTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDFTHLQGGSTLTLESPPGSSPSVQCRSPGKNIQGGKTLVSQLEIQDSG 180
DB 121 LVFGLTANSDFTHLQGGSTLTLESPPGSSPSVQCRSPGKNIQGGKTLVSQLEIQDSG 180
QY 181 TWTCVTLQOKKVEFKIDIVVLAFOKASSIVYKKEGEVFSFPPLAFTVETKLTSGSELMW 240
DB 181 TWTCVTLQOKKVEFKIDIVVLAFOKASSIVYKKEGEVFSFPPLAFTVETKLTSGSELMW 240
QY 241 QAEKASSSKSWITPDLKKNKESVKEKVTODPKLQMGKULPLHLTLPOALPOYAGSGNLTIA 300
DB 241 QAEKASSSKSWITPDLKKNKESVKEKVTODPKLQMGKULPLHLTLPOALPOYAGSGNLTIA 300
QY 301 LEAKTGKLEHVEVNLVYMRATOLQKULTCVWGPSTSPKMLSLKENKAKVSKREKPVWV 360
DB 301 LEAKTGKLEHVEVNLVYMRATOLQKULTCVWGPSTSPKMLSLKENKAKVSKREKPVWV 360
QY 361 LNPBAGMOCCLSDSGOVLLESINIKVLPWTSIPVHADPGEPSCKHTCPPCAPBELL 414
DB 361 LNPBAGMOCCLSDSGOVLLESINIKVLPWTSIPVHADPGEPSCKHTCPPCAPBELL 420
QY 415 GGPVSFLPPPKKDTLMSRTPEVTCVVDVSHEDPEVFNMYVDGVEYHNAKTKPREBO 474
DB 415 GGPVSFLPPPKKDTLMSRTPEVTCVVDVSHEDPEVFNMYVDGVEYHNAKTKPREBO 479
QY 475 YNSTRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSR 534
DB 475 YNSTRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSR 538
QY 535 DELTKNOVSLTCLVNGFVPSDIAVEMESNGOPENNKTTPTVLDSGSEFLLSKLTVDKS 594
DB 535 DELTKNOVSLTCLVNGFVPSDIAVEMESNGOPENNKTTPTVLDSGSEFLLSKLTVDKS 598
QY 595 RMQGNVFSQVMEALHNHYTOKSLSPG 625
DB 595 RMQGNVFSQVMEALHNHYTOKSLSPG 629

```

RESULT 2

```

US-08-477-460B-4
; Sequence 4, Application US/08477460B
; Patent No. 6034223

```

```

; GENERAL INFORMATION:
; APPLICANT: Progenics Pharmaceuticals, Inc.
; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
; TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONJUGATES, AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.24

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,460B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/927,931
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/ATM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 977-9809
; TELETYPE: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: homo sapien
; CELL TYPE: lymphocyte
; US-08-477-460B-4

```

```

Query Match      62.0%; Score 2116; DB 3; Length 530;
Best Local Similarity 68.9%; Pred. No. 7.3e-160;
Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

```

```

QY 1 MNRGVPFRHLVLLVQLALLPAATGKNVVLGKGGDTVELCTASOKKSIOFHMKNNSNOK 60
DB 1 MNRGVPFRHLVLLVQLALLPAATGKNVVLGKGGDTVELCTASOKKSIOFHMKNNSNOK 60
QY 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDTHLLQGSGLTTLTSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSGLTTLTSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANSDTHLLQGSGLTTLTSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSGLTTLTSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIYKKKEGVEFSPLAFVETLTSSELMMW 240
DB 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIYKKKEGVEFSPLAFVETLTSSELMMW 240
QY 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIYKKKEGVEFSPLAFVETLTSSELMMW 240
DB 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIYKKKEGVEFSPLAFVETLTSSELMMW 240
QY 241 QAERASSKSWITFDLKNKEVSVKRVTPQPKLQMGKPLPLHLTPQALPYAG--SGNL 297
DB 241 QAERASSKSWITFDLKNKEVSVKRVTPQPKLQMGKPLPLHLTPQALPYAG--SGNL 297
QY 217 -----PCSRSTSESTALGCLVKDYFPEPVTVSNMNSGALTSQVH 255
DB 217 -----PCSRSTSESTALGCLVKDYFPEPVTVSNMNSGALTSQVH 255
QY 298 TLAEAKTGKLGHOENVLVNMRATOL-QKNLTCEVWGPTSPRLMLSLKENKAVSRKREK 356
DB 298 TLAEAKTGKLGHOENVLVNMRATOL-QKNLTCEVWGPTSPRLMLSLKENKAVSRKREK 356
QY 256 TFPALVQSSGLYSLSVTVPSNFGTQTYCNV-----DHK 292
DB 256 TFPALVQSSGLYSLSVTVPSNFGTQTYCNV-----DHK 292
QY 357 PAVVLNBPAGMOCCLSDSGVLESNIKVLPTWSTPV-----PCPAPBPCKDKTHCP 411
DB 357 PAVVLNBPAGMOCCLSDSGVLESNIKVLPTWSTPV-----PCPAPBPCKDKTHCP 411
QY 293 P-----SNTKVDKTVERKCCVCEPCPAP----- 317
DB 293 P-----SNTKVDKTVERKCCVCEPCPAP----- 317
QY 412 ELGSPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGVGVHNAKTKPR 471
DB 412 ELGSPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGVGVHNAKTKPR 471
QY 318 --VAGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGVGVHNAKTKPR 375
DB 318 --VAGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGVGVHNAKTKPR 375
QY 472 EEOYNSTYRVSVLTVLHODMLNGKEYKCVSNKALPAPIEKTSKAGQPREQVYTLR 531
DB 472 EEOYNSTYRVSVLTVLHODMLNGKEYKCVSNKALPAPIEKTSKAGQPREQVYTLR 531
QY 376 EEOGNSTFRVSVLTVVHODMLNGKEYKCVSNKALPAPIEKTSKAGQPREQVYTLR 435
DB 376 EEOGNSTFRVSVLTVVHODMLNGKEYKCVSNKALPAPIEKTSKAGQPREQVYTLR 435
QY 532 PSRELTKNOVSLTCLVKGFFPSDIAVEMNSGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
DB 532 PSRELTKNOVSLTCLVKGFFPSDIAVEMNSGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
QY 436 PSREMTKNOVSLTCLVKGFFPSDIAVEMNSGQPENNYKTTTPVLDSDGSFFLYSKLTV 495
DB 436 PSREMTKNOVSLTCLVKGFFPSDIAVEMNSGQPENNYKTTTPVLDSDGSFFLYSKLTV 495
QY 592 DKSRMOQGNVFCGVMEALHNHYTOKSLSPG 625
DB 592 DKSRMOQGNVFCGVMEALHNHYTOKSLSPG 625

```

```

DB 496 DKSRMOQGNVFCGVMEALHNHYTOKSLSPG 529
;
; RESULT 3
; US-08-379-516-4
; Sequence 4, Application US/08379516
; Patent No. 6083478
; GENERAL INFORMATION:
; APPLICANT: Allaway, Graham P.
; TITLE OF INVENTION: No. 6083478-Peptidyl Moiety-Conjugated CD4-Gamma2 and CD4-IgG2
; FILE REFERENCE: 41215-A-PCT-US
; CURRENT APPLICATION NUMBER: US/08/379,516
; EARLIER FILING DATE: 1996-06-10
; EARLIER APPLICATION NUMBER: PCT/US93/07422
; EARLIER FILING DATE: 1993-08-06
; EARLIER APPLICATION NUMBER: 07/927,931
; EARLIER FILING DATE: 1992-08-07
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-08-379-516-4

```

```

Query Match      62.0%; Score 2116; DB 3; Length 530;
Best Local Similarity 68.9%; Pred. No. 7.3e-160;
Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

```

```

QY 1 MNRGVPFRHLVLLVQLALLPAATGKNVVLGKGGDTVELCTASOKKSIOFHMKNNSNOK 60
DB 1 MNRGVPFRHLVLLVQLALLPAATGKNVVLGKGGDTVELCTASOKKSIOFHMKNNSNOK 60
QY 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDTHLLQGSGLTTLTSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSGLTTLTSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANSDTHLLQGSGLTTLTSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSGLTTLTSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIYKKKEGVEFSPLAFVETLTSSELMMW 240
DB 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIYKKKEGVEFSPLAFVETLTSSELMMW 240
QY 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIYKKKEGVEFSPLAFVETLTSSELMMW 240
DB 181 TWCTCTVONOKKVEFKIDIVLAFQKASSIYKKKEGVEFSPLAFVETLTSSELMMW 240
QY 241 QAERASSKSWITFDLKNKEVSVKRVTPQPKLQMGKPLPLHLTPQALPYAG--SGNL 297
DB 241 QAERASSKSWITFDLKNKEVSVKRVTPQPKLQMGKPLPLHLTPQALPYAG--SGNL 297
QY 217 -----PCSRSTSESTALGCLVKDYFPEPVTVSNMNSGALTSQVH 255
DB 217 -----PCSRSTSESTALGCLVKDYFPEPVTVSNMNSGALTSQVH 255
QY 298 TLAEAKTGKLGHOENVLVNMRATOL-QKNLTCEVWGPTSPRLMLSLKENKAVSRKREK 356
DB 298 TLAEAKTGKLGHOENVLVNMRATOL-QKNLTCEVWGPTSPRLMLSLKENKAVSRKREK 356
QY 256 TFPALVQSSGLYSLSVTVPSNFGTQTYCNV-----DHK 292
DB 256 TFPALVQSSGLYSLSVTVPSNFGTQTYCNV-----DHK 292
QY 357 PAVVLNBPAGMOCCLSDSGVLESNIKVLPTWSTPV-----PCPAPBPCKDKTHCP 411
DB 357 PAVVLNBPAGMOCCLSDSGVLESNIKVLPTWSTPV-----PCPAPBPCKDKTHCP 411
QY 293 P-----SNTKVDKTVERKCCVCEPCPAP----- 317
DB 293 P-----SNTKVDKTVERKCCVCEPCPAP----- 317
QY 412 ELGSPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGVGVHNAKTKPR 471
DB 412 ELGSPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGVGVHNAKTKPR 471
QY 318 --VAGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGVGVHNAKTKPR 375
DB 318 --VAGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGVGVHNAKTKPR 375
QY 472 EEOYNSTYRVSVLTVLHODMLNGKEYKCVSNKALPAPIEKTSKAGQPREQVYTLR 531
DB 472 EEOYNSTYRVSVLTVLHODMLNGKEYKCVSNKALPAPIEKTSKAGQPREQVYTLR 531
QY 376 EEOGNSTFRVSVLTVVHODMLNGKEYKCVSNKALPAPIEKTSKAGQPREQVYTLR 435
DB 376 EEOGNSTFRVSVLTVVHODMLNGKEYKCVSNKALPAPIEKTSKAGQPREQVYTLR 435
QY 532 PSRELTKNOVSLTCLVKGFFPSDIAVEMNSGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
DB 532 PSRELTKNOVSLTCLVKGFFPSDIAVEMNSGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
QY 436 PSREMTKNOVSLTCLVKGFFPSDIAVEMNSGQPENNYKTTTPVLDSDGSFFLYSKLTV 495
DB 436 PSREMTKNOVSLTCLVKGFFPSDIAVEMNSGQPENNYKTTTPVLDSDGSFFLYSKLTV 495
QY 592 DKSRMOQGNVFCGVMEALHNHYTOKSLSPG 625
DB 592 DKSRMOQGNVFCGVMEALHNHYTOKSLSPG 625

```

Db 496 DKSRMOQGNVFSQSVHMEALHNHYTOKSLSPG 529

RESULT 4
US-09-329-916-4
Sequence 4, Application US/09329916
Patent No. 6177549

GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/329,916
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,460
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCF/JPM/AJM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-329-916-4

Query Match 62.0%; Score 2116; DB 3; Length 530;
Best Local Similarity 68.9%; Pred. No. 7,3e-160;
Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

QY 1 MNRGVPFRHLLVQLALLPAATQGNVLSGKGDVLTCTASQKKSIFPHWKNNOIK 60
DB 1 MNRGVPFRHLLVQLALLPAATQGNVLSGKGDVLTCTASQKKSIFPHWKNNOIK 60
QY 61 ILNGOGSFLTKGPKSLMDRADSRSLMDQGNFPLITNKLITSDTYICEVEOQKEVOL 120
DB 61 ILNGOGSFLTKGPKSLMDRADSRSLMDQGNFPLITNKLITSDTYICEVEOQKEVOL 120
QY 121 LVFGLTANSDTHLLQGSGLTTLTSPGSSPSVQCRSPKNTIOGKTLISVSOLELQDSG 180
DB 121 LVFGLTANSDTHLLQGSGLTTLTSPGSSPSVQCRSPKNTIOGKTLISVSOLELQDSG 180
QY 181 TWTCTVLOQNKQVEFKIDIVLAFOKASSIVYKKEGSEOVERSPFLATVEKLTGSGELMW 240
DB 181 TWTCTVLOQNKQVEFKIDIVLAFOKASSIVYKKEGSEOVERSPFLATVEKLTGSGELMW 240
QY 216 TWTCTVLOQNKQVEFKIDIVLAFOKASSIVYKKEGSEOVERSPFLATVEKLTGSGELMW 216

QY 241 QAERASSKSWITFDLKNKEVSRYKRYTODPKLQMGKLLPLHLTPQALPQYAG---SGNL 297
DB 217 -----PCSRSTSESTALGCLVXDYFPPEPTVSNNSGALTSQVH 255
QY 298 TLAEAKTGKHOENVLYVMRATQL-QKNLTCEVWGPSTPKMLSLKENKAKVSKREK 356
DB 256 TPEPAVLQSSGLVSLSSVTVPSNFGQTYTCNV-----DHK 292
QY 357 PWWVLNPEAGMOCCLSDSGQVLLBSNIXVLPWSTFV-----PCPAPEKSCDKHTTCP 411
DB 293 P-----SNTKVDKTVKRCVCECPCPAP----- 317
QY 412 ELIGPSVFLFPPKPDITMISRTPEVTCVVDVSHEDPEVKFNMYVDGEVHNAKTKPR 471
DB 318 --VAGPSVFLFPPKPDITMISRTPEVTCVVDVSHEDPEVKFNMYVDGEVHNAKTKPR 375
QY 472 EEOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGPREPOVYTLR 531
DB 376 EEOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGPREPOVYTLR 435
QY 532 PSRDELTKQVSLTCLVKGFPYSDIAYEWESNQPENNYKTPPVLDSDSFLYSKLTIV 591
DB 436 PSRDELTKQVSLTCLVKGFPYSDIAYEWESNQPENNYKTPPVLDSDSFLYSKLTIV 495
QY 592 DKSRMOQGNVFSQSVHMEALHNHYTOKSLSPG 625
DB 496 DKSRMOQGNVFSQSVHMEALHNHYTOKSLSPG 529

RESULT 5
US-08-485-372A-4
Sequence 4, Application US/08485372A
Patent No. 6187748

GENERAL INFORMATION:
APPLICANT: Beaudry, Gary A.
TITLE OF INVENTION: CD4-GAMMA2 CD4-IGG2 CHIMERAS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,372A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,227
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:

ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-485-372A-4

Query Match 62.0%; Score 2116; DB 3; Length 530;
Best Local Similarity 68.9%; Pred. No. 7.3e-160;
Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

```

QY 1 MNRGVPFRLHLLVQLALLPAATGKNGKGVTELTCTASOKSIOFHMNSNOIK 60
DB 1 MNRGVPFRLHLLVQLALLPAATGKNGKGVTELTCTASOKSIOFHMNSNOIK 60
QY 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFLLINKLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFLLINKLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTNSDTHLLQGSLLTLESPPSSPSVQCRSPRGNIOGKTLVSQLELDQSG 180
DB 121 LVFGLTNSDTHLLQGSLLTLESPPSSPSVQCRSPRGNIOGKTLVSQLELDQSG 180
QY 121 LVFGLTNSDTHLLQGSLLTLESPPSSPSVQCRSPRGNIOGKTLVSQLELDQSG 180
DB 121 LVFGLTNSDTHLLQGSLLTLESPPSSPSVQCRSPRGNIOGKTLVSQLELDQSG 180
QY 181 TWCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEQVEESFLATVEKLTSGELMW 240
DB 181 TWCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEQVEESFLATVEKLTSGELMW 240
QY 241 QAERASSKSWITFDLKNKEVSVRVQDPKLGKGLPLHLTLPOALPOYAG--SGNL 297
DB 241 QAERASSKSWITFDLKNKEVSVRVQDPKLGKGLPLHLTLPOALPOYAG--SGNL 297
QY 298 TLAEAKTGKLGHEVNLVVMRATOL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
DB 298 TLAEAKTGKLGHEVNLVVMRATOL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
QY 357 PWWLNPBAGWMOCLSDSQVLLSNIKYLPTWSTPV-----PCPAPEKSCDKHTCP 411
DB 357 PWWLNPBAGWMOCLSDSQVLLSNIKYLPTWSTPV-----PCPAPEKSCDKHTCP 411
QY 412 ELGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYDGVENVNAKTKPR 471
DB 412 ELGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYDGVENVNAKTKPR 471
QY 472 EEOYNSTYRVSVLTIVHODMLNGEKYCKVSNKALPAPIEKTSKAKGQPREPOVYTL 531
DB 472 EEOYNSTYRVSVLTIVHODMLNGEKYCKVSNKALPAPIEKTSKAKGQPREPOVYTL 531
QY 532 PSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
DB 532 PSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
QY 592 DKSRLMOGNVPSGCVMEALHNHYTKSLSPG 625
DB 592 DKSRLMOGNVPSGCVMEALHNHYTKSLSPG 625

```

RESULT 6
US-09-409-006A-4
Sequence 4, Application US/09409006A
Patent No. 6342586
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONJUGATES, AND USES THEREOF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/409,006A
FILING DATE: 29-SEP-1999
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/ALM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-409-006A-4

Query Match 62.0%; Score 2116; DB 4; Length 530;
Best Local Similarity 68.9%; Pred. No. 7.3e-160;
Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

```

QY 1 MNRGVPFRLHLLVQLALLPAATGKNGKGVTELTCTASOKSIOFHMNSNOIK 60
DB 1 MNRGVPFRLHLLVQLALLPAATGKNGKGVTELTCTASOKSIOFHMNSNOIK 60
QY 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFLLINKLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFLLINKLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTNSDTHLLQGSLLTLESPPSSPSVQCRSPRGNIOGKTLVSQLELDQSG 180
DB 121 LVFGLTNSDTHLLQGSLLTLESPPSSPSVQCRSPRGNIOGKTLVSQLELDQSG 180
QY 121 LVFGLTNSDTHLLQGSLLTLESPPSSPSVQCRSPRGNIOGKTLVSQLELDQSG 180
DB 121 LVFGLTNSDTHLLQGSLLTLESPPSSPSVQCRSPRGNIOGKTLVSQLELDQSG 180
QY 181 TWCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEQVEESFLATVEKLTSGELMW 240
DB 181 TWCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEQVEESFLATVEKLTSGELMW 240
QY 241 QAERASSKSWITFDLKNKEVSVRVQDPKLGKGLPLHLTLPOALPOYAG--SGNL 297
DB 241 QAERASSKSWITFDLKNKEVSVRVQDPKLGKGLPLHLTLPOALPOYAG--SGNL 297
QY 298 TLAEAKTGKLGHEVNLVVMRATOL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
DB 298 TLAEAKTGKLGHEVNLVVMRATOL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
QY 357 PWWLNPBAGWMOCLSDSQVLLSNIKYLPTWSTPV-----PCPAPEKSCDKHTCP 411
DB 357 PWWLNPBAGWMOCLSDSQVLLSNIKYLPTWSTPV-----PCPAPEKSCDKHTCP 411
QY 412 ELGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYDGVENVNAKTKPR 471
DB 412 ELGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYDGVENVNAKTKPR 471
QY 472 EEOYNSTYRVSVLTIVHODMLNGEKYCKVSNKALPAPIEKTSKAKGQPREPOVYTL 531
DB 472 EEOYNSTYRVSVLTIVHODMLNGEKYCKVSNKALPAPIEKTSKAKGQPREPOVYTL 531
QY 532 PSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
DB 532 PSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
QY 592 DKSRLMOGNVPSGCVMEALHNHYTKSLSPG 625
DB 592 DKSRLMOGNVPSGCVMEALHNHYTKSLSPG 625

```


Best Local Similarity 68.9%; Pred. No. 7,3e-160; Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

```

Qy 1 MNRGVPRHLLLVQLALLPAATGKRVYLGKGDVETLTCTAOKSIOFHMKNSQIK 60
Db 1 MNRGVPRHLLLVQLALLPAATGKRVYLGKGDVETLTCTAOKSIOFHMKNSQIK 60
Qy 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEEYQL 120
Db 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEEYQL 120
Qy 121 LVFGITANSPTHLLOQGSITLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDQSG 180
Db 121 LVFGITANSPTHLLOQGSITLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDQSG 180
Qy 181 TWCTCTVLOKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSFLAFTVEKLTSSGELMW 240
Db 181 TWCTCTVLOKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSFLAFTVEKLTSSGELMW 240
Qy 241 QAERASSKSMITTDLNKKEVSVRVYQDPKLOMGKKLPLHLTLPOALPOYAG--SGNL 297
Db 241 QAERASSKSMITTDLNKKEVSVRVYQDPKLOMGKKLPLHLTLPOALPOYAG--SGNL 297
Qy 298 TLAEAKTGKLEHVEVNLVVRATOL-QKULTCEVWGPTSPKLMLSKLENKAKVSKREK 356
Db 298 TLAEAKTGKLEHVEVNLVVRATOL-QKULTCEVWGPTSPKLMLSKLENKAKVSKREK 356
Qy 357 PVMWLNPEAGMOCILSDSGVLLSNIKVLPTWSTPV---PCPAPEKSCDKTHTCP 411
Db 357 PVMWLNPEAGMOCILSDSGVLLSNIKVLPTWSTPV---PCPAPEKSCDKTHTCP 411
Qy 412 ELGGPSVFLPPKPKOTLMTSRPEVVCVVVDVSHEDPEVKFMVYDGEVHNAKTKPR 471
Db 412 ELGGPSVFLPPKPKOTLMTSRPEVVCVVVDVSHEDPEVKFMVYDGEVHNAKTKPR 471
Qy 472 EEOYNSTRVYVSVLTLHODWLNKKEYKCKVSNKALPAPIEKITSKAGQPREPOVYTLF 531
Db 472 EEOYNSTRVYVSVLTLHODWLNKKEYKCKVSNKALPAPIEKITSKAGQPREPOVYTLF 531
Qy 532 PSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPYLDSGSEFFLYSKLTV 591
Db 532 PSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPYLDSGSEFFLYSKLTV 591
Qy 592 DKSRMOQGNVPSGVMEALHNHYTKSLSPG 625
Db 592 DKSRMOQGNVPSGVMEALHNHYTKSLSPG 625
Qy 496 DKSRMOQGNVPSGVMEALHNHYTKSLSPG 529
Db 496 DKSRMOQGNVPSGVMEALHNHYTKSLSPG 529

RESULT 9
US-08-477-460B-2
; Sequence 2, Application US/08477460B
; Patent No. 6034223
; GENERAL INFORMATION:
; APPLICANT: Progenics Pharmaceuticals, Inc.
; TITLE OF INVENTION: CD4-GAMMA2 AND CD4-1G62 IMMUNOCONJUGATES, AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,460B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530

```

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/927,931

FILING DATE: 07-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 977-9805

TELEX: 422523 COOP UI

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 432 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: homo sapien

CELL TYPE: lymphocyte

US-08-477-460B-2

Query Match

Best Local Similarity 60.3%; Score 2059; DB 3; Length 432;

Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

```

Qy 1 MNRGVPRHLLLVQLALLPAATGKRVYLGKGDVETLTCTAOKSIOFHMKNSQIK 60
Db 1 MNRGVPRHLLLVQLALLPAATGKRVYLGKGDVETLTCTAOKSIOFHMKNSQIK 60
Qy 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEEYQL 120
Db 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKEEYQL 120
Qy 121 LVFGITANSPTHLLOQGSITLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDQSG 180
Db 121 LVFGITANSPTHLLOQGSITLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDQSG 180
Qy 181 TWCTCTVLOKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSFLAFTVEKLTSSGELMW 240
Db 181 TWCTCTVLOKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSFLAFTVEKLTSSGELMW 240
Qy 241 QAERASSKSMITTDLNKKEVSVRVYQDPKLOMGKKLPLHLTLPOALPOYAG--SGNL 300
Db 241 QAERASSKSMITTDLNKKEVSVRVYQDPKLOMGKKLPLHLTLPOALPOYAG--SGNL 300
Qy 301 LEAKTGKLEHVEVNLVVRATOL-QKULTCEVWGPTSPKLMLSKLENKAKVSKREKPVW 360
Db 301 LEAKTGKLEHVEVNLVVRATOL-QKULTCEVWGPTSPKLMLSKLENKAKVSKREKPVW 360
Qy 361 LNPBAGMOCILSDSGVLLSNIKVLPTWSTPVPCPAPEKSCDKTHTCPDELLGSPVF 420
Db 361 LNPBAGMOCILSDSGVLLSNIKVLPTWSTPVPCPAPEKSCDKTHTCPDELLGSPVF 420
Qy 421 LFPKPKOTLMTSRPEVVCVVVDVSHEDPEVKFMVYDGEVHNATKPREQYNSYR 480
Db 421 LFPKPKOTLMTSRPEVVCVVVDVSHEDPEVKFMVYDGEVHNATKPREQYNSYR 480
Qy 481 VVSVLTLHODWLNKKEYKCKVSNKALPAPIEKITSKAGQPREPOVYTLPPSRDELTKN 540
Db 481 VVSVLTLHODWLNKKEYKCKVSNKALPAPIEKITSKAGQPREPOVYTLPPSRDELTKN 540
Qy 541 QVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPYLDSGSEFFLYSKLTVDSKMOQGN 600
Db 541 QVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPYLDSGSEFFLYSKLTVDSKMOQGN 600
Qy 601 VFCSSVMHEALHNHYTKSLSPG 625
Db 601 VFCSSVMHEALHNHYTKSLSPG 625
Qy 407 VFCSSVMHEALHNHYTKSLSPG 431
Db 407 VFCSSVMHEALHNHYTKSLSPG 431

RESULT 10

```

```
US-08-379-516-2
; Sequence 2, Application US/08379516
; Patent No. 6083478
; GENERAL INFORMATION:
; APPLICANT: Allaway, Graham P.
; TITLE OF INVENTION: No. 6083478-Peptidyl Molecly-Confugated CD4-Gamma2 and CD4-IgG2
; TITLE OF INVENTION: Immunocnjugates and Uses Thereof
; FILE REFERENCE: 41215-A-PCT-US
; CURRENT APPLICATION NUMBER: US/08/379, 516
; CURRENT FILING DATE: 1996-06-10
; EARLIER APPLICATION NUMBER: PCT/US93/07422
; EARLIER FILING DATE: 1993-08-06
; EARLIER APPLICATION NUMBER: 07/927, 931
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; LENGTH: 432
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-379-516-2

Query Match      60.3%; Score 2059; DB 3; Length 432;
Best Local Similarity 65.6%; Pred. No. 1.8e-155;
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

QY 1 MNRGVPRHLLVLQALPAATQGNKVLGKKGDVETVCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPRHLLVLQALPAATQGNKVLGKKGDVETVCTASQKKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKIEDSDTYICEVEDQKEEVL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKIEDSDTYICEVEDQKEEVL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLELODSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLELODSG 180
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLELODSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLELODSG 180
QY 181 TWTCVLOQOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWTCVLOQOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
QY 181 TWTCVLOQOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWTCVLOQOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
QY 241 QAEBASSKSWITFDLKNKEVSVKRVYQDPKLOMGKCLPLHLTLPLQALPQYAGSGNLTLLA 300
DB 241 QAEBASSKSWITFDLKNKEVSVKRVYQDPKLOMGKCLPLHLTLPLQALPQYAGSGNLTLLA 300
QY 211 ----- 210
DB 211 ----- 210
QY 301 LEAKTGKLEQEVNLVVRATQLOQNLTCFVWGPTSPKMLMLSKLENKAKVSKREKPVWV 360
DB 301 LEAKTGKLEQEVNLVVRATQLOQNLTCFVWGPTSPKMLMLSKLENKAKVSKREKPVWV 360
QY 211 ----- 210
DB 211 ----- 210
QY 361 LNPBAGMOCLLSDSGVLLBSNLIKVLPTWSTVPYCPAPBPBKSCDKHTHCPBLLGSPSVF 420
DB 361 LNPBAGMOCLLSDSGVLLBSNLIKVLPTWSTVPYCPAPBPBKSCDKHTHCPBLLGSPSVF 420
QY 211 -----EC-----PCCPAPP-----VAGSPVF 226
DB 211 -----EC-----PCCPAPP-----VAGSPVF 226
QY 421 LFPKPKDPTLMISTPREVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSYTR 480
DB 421 LFPKPKDPTLMISTPREVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSYTR 480
QY 227 LFPKPKDPTLMISTPREVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSYTR 286
DB 227 LFPKPKDPTLMISTPREVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSYTR 286
QY 481 VVSVLTIVHODMLNGEKYCKCKVSNKALPAPLEKTIISKAKGQPREPQVYTLPPSRDELTKN 540
DB 481 VVSVLTIVHODMLNGEKYCKCKVSNKALPAPLEKTIISKAKGQPREPQVYTLPPSRDELTKN 540
QY 287 VVSVLTIVHODMLNGEKYCKCKVSNKALPAPLEKTIISKAKGQPREPQVYTLPPSRDELTKN 346
DB 287 VVSVLTIVHODMLNGEKYCKCKVSNKALPAPLEKTIISKAKGQPREPQVYTLPPSRDELTKN 346
QY 541 QVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPRVLDSDGSPFLYSKLTVDKSRMQGN 600
DB 541 QVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPRVLDSDGSPFLYSKLTVDKSRMQGN 600
QY 347 QVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPRVLDSDGSPFLYSKLTVDKSRMQGN 406
DB 347 QVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPRVLDSDGSPFLYSKLTVDKSRMQGN 406
QY 601 VFSCSVNHEALHNHTYTKSLSLSPG 625
DB 601 VFSCSVNHEALHNHTYTKSLSLSPG 625
QY 407 VFSCSVNHEALHNHTYTKSLSLSPG 431
DB 407 VFSCSVNHEALHNHTYTKSLSLSPG 431
```

```
US-09-329-916-2
; Sequence 2, Application US/09329916
; Patent No. 617549
; GENERAL INFORMATION:
; APPLICANT: Progenics Pharmaceuticals, Inc.
; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
; TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IgG2 IMMUNOCNUGATES, AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/329, 916
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/477,460
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 07/927, 931
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/AJM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 977-9809
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; CELL TYPE: lymphocyte
US-09-329-916-2

Query Match      60.3%; Score 2059; DB 3; Length 432;
Best Local Similarity 65.6%; Pred. No. 1.8e-155;
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

QY 1 MNRGVPRHLLVLQALPAATQGNKVLGKKGDVETVCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPRHLLVLQALPAATQGNKVLGKKGDVETVCTASQKKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKIEDSDTYICEVEDQKEEVL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKIEDSDTYICEVEDQKEEVL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLELODSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLELODSG 180
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLELODSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVQLELODSG 180
QY 181 TWTCVLOQOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWTCVLOQOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
QY 181 TWTCVLOQOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWTCVLOQOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
QY 241 QAEBASSKSWITFDLKNKEVSVKRVYQDPKLOMGKCLPLHLTLPLQALPQYAGSGNLTLLA 300
DB 241 QAEBASSKSWITFDLKNKEVSVKRVYQDPKLOMGKCLPLHLTLPLQALPQYAGSGNLTLLA 300
QY 211 ----- 210
DB 211 ----- 210
```

```

Qy 301 LEAKTKLHGEVNLVWRATQLOKNLTCEVWGPTSPKLTLSLKLENKEAKVSKREKPYWV 360
Db 211 ----- 210
Qy 361 LNPEAGMOCLLSDSGVLLSNIKVLPTWSTPVPCPAPBPXSCDKHTCPBELLGSPVF 420
Db 211 -----EC-----PPCPAP-----VAGPSVF 226
Qy 421 LPPPKDGLMISRTPEVTCVVDVSHDEPVEKFMVYDGVGVNNAKTKPREBOYNSYR 480
Db 227 LPPPKDGLMISRTPEVTCVVDVSHDEPVEVQFMVYDGVGVNNAKTKPREBOYNSYR 286
Qy 481 VVSULTVHODMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSDELTKN 540
Db 287 VVSULTVHODMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSDELTKN 346
Qy 541 QVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFLYSKLTVDKSRMWOQN 600
Db 347 QVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFLYSKLTVDKSRMWOQN 406
Qy 601 VFSCSVMEALHNHYTOKSLSPG 625
Db 407 VFSCSVMEALHNHYTOKSLSPG 431

```

RESULT 12

US-08-485-372A-2

Sequence 2, Application US/08485372A

Patent No. 6187748

GENERAL INFORMATION:

APPLICANT: Beaudry, Gary A.

APPLICANT: Maddon, Paul J.

TITLE OF INVENTION: CD4-GAMMA2 CD4-IGG2 CHIMERAS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,372A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,227

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 37690-II-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 432 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: homo sapien

CELL TYPE: lymphocyte

US-08-485-372A-2

Query Match 60.3%; Score 2059; DB 3; Length 432;
 Best Local Similarity 65.6%; Pred. No. 1,8e-155;
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

```

Qy 1 MNRGVPFRHLILVQLALLPAATQGNKVLGKGDVLTCTASQKKSIOFHMNSQIK 60
Db 1 MNRGVPFRHLILVQLALLPAATQGNKVLGKGDVLTCTASQKKSIOFHMNSQIK 60
Qy 61 ILGNQGSFLTKGPEKNDNRDSDRSRLMDQGFLLIKNLKEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGPEKNDNRDSDRSRLMDQGFLLIKNLKEDSDTYICEVEDQKEEVOL 120
Qy 121 LVFGLTANSPTHLLOQSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSPTHLLOQSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy 181 TWCTCTVLQNGKVFYKIDIVLAFQKASSIVYKKEGQVEFSFPLAFTVEXLTGSGELMW 240
Db 181 TWCTCTVLQNGKVFYKIDIVLAFERKCV----- 210
Qy 241 QAERASSKSKWITFDLKNKEVSVRVQDRKLOMGKTLPLHLTLPOALPOYAGSGNLTLA 300
Db 211 ----- 210
Qy 301 LEAKTKLHGEVNLVWRATQLOKNLTCEVWGPTSPKLTLSLKLENKEAKVSKREKPYWV 360
Db 211 ----- 210
Qy 361 LNPEAGMOCLLSDSGVLLSNIKVLPTWSTPVPCPAPBPXSCDKHTCPBELLGSPVF 420
Db 211 -----EC-----PPCPAP-----VAGPSVF 226
Qy 421 LPPPKDGLMISRTPEVTCVVDVSHDEPVEKFMVYDGVGVNNAKTKPREBOYNSYR 480
Db 227 LPPPKDGLMISRTPEVTCVVDVSHDEPVEVQFMVYDGVGVNNAKTKPREBOYNSYR 286
Qy 481 VVSULTVHODMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSDELTKN 540
Db 287 VVSULTVHODMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSDELTKN 346
Qy 541 QVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFLYSKLTVDKSRMWOQN 600
Db 347 QVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFLYSKLTVDKSRMWOQN 406
Qy 601 VFSCSVMEALHNHYTOKSLSPG 625
Db 407 VFSCSVMEALHNHYTOKSLSPG 431

```

RESULT 13

US-09-409-006A-2

Sequence 2, Application US/09409006A

Patent No. 6342586

GENERAL INFORMATION:

APPLICANT: Progenics Pharmaceuticals, Inc.

TITLE OF INVENTION: NON-PEPTIDYL NOIETY-CONJUGATED

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/409,006A

FILING DATE: 29-SEP-1999

```

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AM
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-409-006A-2

```

```

Query Match      60.3%; Score 2059; DB 4; Length 432;
Best Local Similarity 65.6%; Pred. No. 1.8e-155;
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

```

```

QY 1 MNRGVPRHLLVLTQALPAATQGNKVVLGKKGDVLTCTASOKKSIQFHKNKSNQIK 60
DB 1 MNRGVPRHLLVLTQALPAATQGNKVVLGKKGDVLTCTASOKKSIQFHKNKSNQIK 60
QY 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCIVLQNKQKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWTCIVLQNKQKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
QY 241 QABASSSKSWITFDLKNKEVSVKRVTDPKLQMGKCLPLHLTPQALPOYAGSGNLTIA 300
DB 241 QABASSSKSWITFDLKNKEVSVKRVTDPKLQMGKCLPLHLTPQALPOYAGSGNLTIA 300
QY 301 LEAKTGKLEHVNVLVVRATQLOKNLTCEVWGFTSPKMLSLKLENKAIVSKREKPVWV 360
DB 301 LEAKTGKLEHVNVLVVRATQLOKNLTCEVWGFTSPKMLSLKLENKAIVSKREKPVWV 360
QY 421 LFPKPKDITLMISTPEVTCVVDVSHEDPEYKRNWVVDGEVHNATKREQYNSTYR 480
DB 421 LFPKPKDITLMISTPEVTCVVDVSHEDPEYKRNWVVDGEVHNATKREQYNSTYR 480
QY 481 VWSVLTAVHODMLNGEKYCKVSNKALPAPIEKTISAKGQPREPOVYTLPPSRDELTKN 540
DB 481 VWSVLTAVHODMLNGEKYCKVSNKALPAPIEKTISAKGQPREPOVYTLPPSRDELTKN 540
QY 541 QVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFELYSKLTVDKSRMQGN 600
DB 541 QVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFELYSKLTVDKSRMQGN 600
QY 601 VFGCSVMEHALHNYTQKSLSLSPG 625
DB 601 VFGCSVMEHALHNYTQKSLSLSPG 625
QY 407 VFCSVMEHALHNYTQKSLSLSPG 431
DB 407 VFCSVMEHALHNYTQKSLSLSPG 431

```

```

RESULT 14
US-08-484-681-2
Sequence 2, Application US/08484681
Patent No. 6451313
GENERAL INFORMATION:
APPLICANT: Beaudry, Gary A.
APPLICANT: Maddon, Paul J.
TITLE OF INVENTION: CD4-GAMMA2 CD4-15G2 CHIMERAS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-484-681-2

```

```

Query Match      60.3%; Score 2059; DB 4; Length 432;
Best Local Similarity 65.6%; Pred. No. 1.8e-155;
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

```

```

QY 1 MNRGVPRHLLVLTQALPAATQGNKVVLGKKGDVLTCTASOKKSIQFHKNKSNQIK 60
DB 1 MNRGVPRHLLVLTQALPAATQGNKVVLGKKGDVLTCTASOKKSIQFHKNKSNQIK 60
QY 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCIVLQNKQKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWTCIVLQNKQKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
QY 241 QABASSSKSWITFDLKNKEVSVKRVTDPKLQMGKCLPLHLTPQALPOYAGSGNLTIA 300
DB 241 QABASSSKSWITFDLKNKEVSVKRVTDPKLQMGKCLPLHLTPQALPOYAGSGNLTIA 300
QY 301 LEAKTGKLEHVNVLVVRATQLOKNLTCEVWGFTSPKMLSLKLENKAIVSKREKPVWV 360
DB 301 LEAKTGKLEHVNVLVVRATQLOKNLTCEVWGFTSPKMLSLKLENKAIVSKREKPVWV 360
QY 407 VFCSVMEHALHNYTQKSLSLSPG 431
DB 407 VFCSVMEHALHNYTQKSLSLSPG 431

```

Qy 361 LNPEAGMOCLLSDSGVLLSNIKVLPTWSTPVPCBPAPKSCDKHTTCPELLGSPSVF 420
Db 211 -----EC-----PCCAPAP-----VAGPSVF 226
Qy 421 LPPPKXTLMTISRTPEVTCVVDVSHDEPVKFNWYVDGVEVNAKTKPREEOYNSTYR 480
Db 227 LPPPKXTLMTISRTPEVTCVVDVSHDEPVQFNMVYDGVENAKTKPREEOFNSTFR 286
Qy 481 VVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKOPREPOVYTLPPSRDELTKN 540
Db 287 VVSVLTVHODMLNGKEYKCKVSNKGLPAPIEKTISKTKGPREPOVYTLPPSRDEMTKN 346
Qy 541 QVSLTCLVKGFPSPDIIVLAEVWESNGQPENNYKTPPVLDSDGSFLLYSKLTVDKSRMOQGN 600
Db 347 QVSLTCLVKGFPSPDIIVLAEVWESNGQPENNYKTPPVLDSDGSFLLYSKLTVDKSRMOQGN 406
Qy 601 VFSCSVMEALHNHYTKSLSPG 625
Db 407 VFSCSVMEALHNHYTKSLSPG 431

RESULT 15

PCT-US93-07422-2
; Sequence 2, Application PC/TUS9307422
; GENERAL INFORMATION:
; APPLICANT: Progenics Pharmaceuticals, Inc.
; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
; TITLE OF INVENTION: CD4-GMM2 AND CD4-1GG2 IMMUNOCONJUGATES, AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07422
; FILING DATE: 19930806
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/927,931
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/ADM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 977-9809
; TELEX: 422523 COOP UT
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: homo sapien
; CELL TYPE: lymphocyte
PCT-US93-07422-2

Query Match 60.3%; Score 2059; DB 5; Length 432;
Beet Local Similarity 65.6%; Pred. No. 1.8e-155;
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;
Qy 1 MNRGVPFRHLLLVQLALLPAATQGNKVVLGKKGDTVELTCTASOKSIOFHWKNSNQIK 60

Db 1 MNRGVPFRHLLLVQLALLPAATQGNKVVLGKKGDTVELTCTASOKSIOFHWKNSNQIK 60
Qy 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
Db 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
Qy 121 LVFGLTANSPDTHLQSGSLTLTLESPPGSSPSVOCSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSPDTHLQSGSLTLTLESPPGSSPSVOCSPRGKNIQGGKTLVSQLELDQSG 180
Qy 181 TWCTTVLQONQKVEFKIDIVLAEQKASSIVYKKEGQVEFSFPLAEFTVEKLTGSGELMW 240
Db 181 TWCTTVLQONQKVEFKIDIVLAEKRCV----- 210
Qy 241 QAEKASSKSWITDLDKNEKVSVRVTQDPKLGKCLPLHLTLPOLPOYAGSGNLTLA 300
Db 211 ----- 210
Qy 301 LEAKTGKHQEVNVLVWRATQLOKNLTCVWGPSPKMLSLKENKAQVSKREKPVWV 360
Db 211 ----- 210
Qy 361 LNPEAGMOCLLSDSGVLLSNIKVLPTWSTPVPCBPAPKSCDKHTTCPELLGSPSVF 420
Db 211 -----EC-----PCCAPAP-----VAGPSVF 226
Qy 421 LPPPKXTLMTISRTPEVTCVVDVSHDEPVKFNWYVDGVEVNAKTKPREEOYNSTYR 480
Db 227 LPPPKXTLMTISRTPEVTCVVDVSHDEPVQFNMVYDGVENAKTKPREEOFNSTFR 286
Qy 481 VVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKOPREPOVYTLPPSRDELTKN 540
Db 287 VVSVLTVHODMLNGKEYKCKVSNKGLPAPIEKTISKTKGPREPOVYTLPPSRDEMTKN 346
Qy 541 QVSLTCLVKGFPSPDIIVLAEVWESNGQPENNYKTPPVLDSDGSFLLYSKLTVDKSRMOQGN 600
Db 347 QVSLTCLVKGFPSPDIIVLAEVWESNGQPENNYKTPPVLDSDGSFLLYSKLTVDKSRMOQGN 406
Qy 601 VFSCSVMEALHNHYTKSLSPG 625
Db 407 VFSCSVMEALHNHYTKSLSPG 431

RESULT 16

US-08-417-495-6
; Sequence 6, Application US/08417495
; Patent No. 5843728
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Redirection of Cellular Immunity by Chimerae
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,495
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,866
; FILING DATE:
; APPLICATION NUMBER: US/07/847,566

FILING DATE:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-417-495-6

Query Match 59.7%; Score 2036.5; DB 2; Length 532;
Best Local Similarity 98.3%; Pred. No. 1.5e-153;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPFRLLLVLTALPAPATQGNKVVGGKGDVTELTCTASQKSIQFHKNSNQIK 60
DB 1 MNRGVPFRLLLVLTALPAPATQGNKVVGGKGDVTELTCTASQKSIQFHKNSNQIK 60
QY 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNFPLIIKNKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNFPLIIKNKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTRANSDTHLLOQOSLTLTLESPPGSSPVQCSPPGKNIQGGKTLVSQLELODSG 180
DB 121 LVFGLTRANSDTHLLOQOSLTLTLESPPGSSPVQCSPPGKNIQGGKTLVSQLELODSG 180
QY 181 TWCTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMW 240
DB 181 TWCTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLNKKEVSVKRVYQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
DB 241 QAERASSSKSWITFDLNKKEVSVKRVYQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
QY 301 LEAKTGKLEHGVNLVVMRATQLOKNTCEVWGPTSPKMLSLKLENKAKVSKREKPVWV 360
DB 301 LEAKTGKLEHGVNLVVMRATQLOKNTCEVWGPTSPKMLSLKLENKAKVSKREKPVWV 360
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPPEKSC 404
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPPEKSC 404

RESULT 17
US-08-284-391B-6
Sequence 6, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanue, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-6

Query Match 59.7%; Score 2036.5; DB 2; Length 532;
Best Local Similarity 98.3%; Pred. No. 1.5e-153;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPFRLLLVLTALPAPATQGNKVVGGKGDVTELTCTASQKSIQFHKNSNQIK 60
DB 1 MNRGVPFRLLLVLTALPAPATQGNKVVGGKGDVTELTCTASQKSIQFHKNSNQIK 60
QY 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNFPLIIKNKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNFPLIIKNKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTRANSDTHLLOQOSLTLTLESPPGSSPVQCSPPGKNIQGGKTLVSQLELODSG 180
DB 121 LVFGLTRANSDTHLLOQOSLTLTLESPPGSSPVQCSPPGKNIQGGKTLVSQLELODSG 180
QY 181 TWCTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMW 240
DB 181 TWCTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLNKKEVSVKRVYQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
DB 241 QAERASSSKSWITFDLNKKEVSVKRVYQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
QY 301 LEAKTGKLEHGVNLVVMRATQLOKNTCEVWGPTSPKMLSLKLENKAKVSKREKPVWV 360
DB 301 LEAKTGKLEHGVNLVVMRATQLOKNTCEVWGPTSPKMLSLKLENKAKVSKREKPVWV 360
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPPEKSC 404
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPPEKSC 404

RESULT 18
US-09-218-950-6
Sequence 6, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles

```
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-6

Query Match      59.7%; Score 2036.5; DB 3; Length 532;
Best Local Similarity 98.3%; Pred. No. 1.5e-153;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;
```

```
361 LNPERAGWQCLSDSGVLTLESNIKVLPTWSTPYPCAPERKSC 404
|||||
361 LNPERAGWQCLSDSGVLTLESNIKVLPTWSTPYPHA---DKLKC 401

RESULT 19
PCT-US92-01785-6
Sequence 6, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Mordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-01785-6

Query Match      59.7%; Score 2036.5; DB 5; Length 532;
Best Local Similarity 98.3%; Pred. No. 1.5e-153;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;
```

Qy	181	TWCTVQLONCKKVEFKIDIVYLAFOKASSIVYKKBEQVSEFPLAFTVEKLTGSGELMW	240
Db	181	TWCTVQLONCKKVEFKIDIVLAFOKASSIVYKKEGVSEFPLAFTVEKLTGSGELMW	240
Qy	241	QARRASSSKSWITFDLKNEKESVKRYTODPKLOMGKKLPHLTLPOLPOAGSGLTTLA	300
Db	241	QARRASSSKSWITFDLKNEKESVKRYTODPKLOMGKKLPHLTLPOLPOAGSGLTTLA	300
Qy	301	LEAKTGKLHOENVLVMMRATOLQKNLTCEVMGTPSPKMLSLKENKAAYSREKPVWV	360
Db	301	LEAKTGKLHOENVLVMMRATOLQKNLTCEVMGTPSPKMLSLKENKAAYSREKPVWV	360
Qy	361	LNEAGMWOCLLSDSGVLLSNIKYLPMTSTVPCCAPBPKSC	404
Db	361	LNEAGMWOCLLSDSGVLLSNIKYLPMTSTVHA--DPKLC	401

RESULT 21
US-08-417-495-4
Sequence 4, Application US/08417495
Patent No. 5843728
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by Chimerae
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,495
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/203,866
FILING DATE:
APPLICATION NUMBER: US/07/847,566
FILING DATE:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-417-495-4

Query Match	59.7%;	Score 2036.5;	DB 2;	Length 575;
Best Local Similarity	98.3%;	Pred. No. 1,7e-153;		
Matches 397;	Conservative 1;	Mismatches 3;	Indels 3;	Gaps 1;

```

1 MNRGVPRRLILVLQALLPAATGNKVYLGGKDTVELTCTASOKSIGPHMKNSNOIK 60
|||||

```


Db 1 MNRGVPRHLLVLIQALLPAATQGNKRVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60
QY 61 ILGNQGSFLTKGPKSLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGPKSLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLLOQGSLLTTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
Db 121 LVFGLTANSPTHLLOQGSLLTTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
QY 181 TWCTVLOQKQKVEFKIDIVVLAFOKASSIYKKEGQVEFSPLAFTVEKLTGSGELMW 240
Db 181 TWCTVLOQKQKVEFKIDIVVLAFOKASSIYKKEGQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQMGKKLPLHLTLPOLPOYAGSGNLTLA 300
Db 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQMGKKLPLHLTLPOLPOYAGSGNLTLA 300
QY 301 LEAKTGKLOHEVNLVWMRATOLQKLTCEVWGPTSPKMLSLKLENKAKYSKEKPYWV 360
Db 301 LEAKTGKLOHEVNLVWMRATOLQKLTCEVWGPTSPKMLSLKLENKAKYSKEKPYWV 360
QY 361 LNPAGMWQCLSDSGVLTLESNIKVLPTWSTPVPCPAPEBKSC 404
Db 361 LNPAGMWQCLSDSGVLTLESNIKVLPTWSTPVPHA--DBKLC 401

RESULT 22
US-08-284-391B-4
; Sequence 4, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR - BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391B
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids

#: TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-284-391B-4

Query Match 59.3%; Score 2036.5; DB 2; Length 575;
Best Local Similarity 98.3%; Pred. No. 1.7e-153;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPRHLLVLIQALLPAATQGNKRVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60
Db 1 MNRGVPRHLLVLIQALLPAATQGNKRVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60
QY 61 ILGNQGSFLTKGPKSLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGPKSLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLLOQGSLLTTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
Db 121 LVFGLTANSPTHLLOQGSLLTTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
QY 181 TWCTVLOQKQKVEFKIDIVVLAFOKASSIYKKEGQVEFSPLAFTVEKLTGSGELMW 240
Db 181 TWCTVLOQKQKVEFKIDIVVLAFOKASSIYKKEGQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQMGKKLPLHLTLPOLPOYAGSGNLTLA 300
Db 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQMGKKLPLHLTLPOLPOYAGSGNLTLA 300
QY 301 LEAKTGKLOHEVNLVWMRATOLQKLTCEVWGPTSPKMLSLKLENKAKYSKEKPYWV 360
Db 301 LEAKTGKLOHEVNLVWMRATOLQKLTCEVWGPTSPKMLSLKLENKAKYSKEKPYWV 360
QY 361 LNPAGMWQCLSDSGVLTLESNIKVLPTWSTPVPCPAPEBKSC 404
Db 361 LNPAGMWQCLSDSGVLTLESNIKVLPTWSTPVPHA--DBKLC 401

RESULT 23
US-09-218-950-4
; Sequence 4, Application US/09218950
; Patent No. 6284240
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR - BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/218,950
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566

```

; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Eilding, Karen L.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
;
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-218-950-4

```

```

Query Match          59.7%; Score 2036.5; DB 3; Length 575;
Best Local Similarity 98.3%; Pred. No. 1,7e-153;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

```

```

QY 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKKGDVTELTCTASQKKSIOFHMKNSNOIK 60
DB 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKKGDVTELTCTASQKKSIOFHMKNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICVEBQKEEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICVEBQKEEVOL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLOQKQVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLOQKQVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
DB 241 QAERASSSSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
QY 301 LEAKTGKLEHVEVNLVYMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
DB 301 LEAKTGKLEHVEVNLVYMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
QY 361 LNPBAGMWOCCLSDSGQVLLSNIKYLPPTWSTVPPCAPPEPKSC 404
DB 361 LNPBAGMWOCCLSDSGQVLLSNIKYLPPTWSTVPVHA---DPKLC 401

```

```

RESULT 24
PCT-US92-01785-4
; Sequence 4, Application PC/TUS9201785
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/01785
; FILING DATE: 19920306
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/665,961
; FILING DATE: March 7, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/119002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
;
; TELETYPE: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-01785-4

```

```

Query Match          59.7%; Score 2036.5; DB 5; Length 575;
Best Local Similarity 98.3%; Pred. No. 1,7e-153;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

```

```

QY 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKKGDVTELTCTASQKKSIOFHMKNSNOIK 60
DB 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKKGDVTELTCTASQKKSIOFHMKNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICVEBQKEEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICVEBQKEEVOL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLOQKQVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLOQKQVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
DB 241 QAERASSSSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
QY 301 LEAKTGKLEHVEVNLVYMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
DB 301 LEAKTGKLEHVEVNLVYMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
QY 361 LNPBAGMWOCCLSDSGQVLLSNIKYLPPTWSTVPPCAPPEPKSC 404
DB 361 LNPBAGMWOCCLSDSGQVLLSNIKYLPPTWSTVPVHA---DPKLC 401

```

```

RESULT 25
PCT-US95-00454-4
; Sequence 4, Application PC/TUS9500454
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 555X
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 SOFTWARE: Wordperfect (Version 5.0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/00454
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/847,566
 FILING DATE: March 6, 1992
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/665,961
 FILING DATE: March 7, 1991
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00786/247001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 575 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-00454-4

Query Match 59.7%; Score 2036.5; DB 5; Length 575;
 Best Local Similarity 98.3%; Pred. No. 1,7e-153;

Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGDVVELCTASQKSIQPHMNSNQIK 60
 DB 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGDVVELCTASQKSIQPHMNSNQIK 60
 QY 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNFPIIKNLKIEDSDTYICEVEDQKEVOL 120
 DB 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNFPIIKNLKIEDSDTYICEVEDQKEVOL 120
 QY 121 LVFGLTANSPTHLIQGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
 DB 121 LVFGLTANSPTHLIQGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
 QY 121 LVFGLTANSPTHLIQGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
 DB 121 LVFGLTANSPTHLIQGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
 QY 181 TWCTVLONOKKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240
 DB 181 TWCTVLONOKKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240
 QY 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLQMGKKLPLHLTLPQALPOYAGSGNLTIA 300
 DB 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLQMGKKLPLHLTLPQALPOYAGSGNLTIA 300
 QY 301 LEATGTGLHGEVNLVVMRATQLOKNTLCEWGPSPKLMSTLKENKEAVSKREKPVWV 360
 DB 301 LEATGTGLHGEVNLVVMRATQLOKNTLCEWGPSPKLMSTLKENKEAVSKREKPVWV 360
 QY 361 LNPAGMOCCLSDSGOVLLESNIKVLPTWSTVPAPAPKSC 404
 DB 361 LNPAGMOCCLSDSGOVLLESNIKVLPTWSTVPAPAPKSC 404

RESULT 26
 US-08-417-495-5
 Sequence 5, Application US/06417495
 Patent No. 5843728
 GENERAL INFORMATION:
 APPLICANT: Seed, Brian et al.
 TITLE OF INVENTION: Redirection of Cellular Immunity by
 TITLE OF INVENTION: Receptor Chimerae

NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or 555X
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 SOFTWARE: Wordperfect (Version 5.0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/417,495
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/203,866
 FILING DATE:
 APPLICATION NUMBER: US/07/847,566
 FILING DATE:
 APPLICATION NUMBER: 07/665,961
 FILING DATE: March 7, 1991
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 00786/119002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 462 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-417-495-5

Query Match 59.5%; Score 2032.5; DB 2; Length 462;
 Best Local Similarity 98.0%; Pred. No. 2.6e-153;

Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGDVVELCTASQKSIQPHMNSNQIK 60
 DB 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGDVVELCTASQKSIQPHMNSNQIK 60
 QY 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNFPIIKNLKIEDSDTYICEVEDQKEVOL 120
 DB 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNFPIIKNLKIEDSDTYICEVEDQKEVOL 120
 QY 121 LVFGLTANSPTHLIQGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
 DB 121 LVFGLTANSPTHLIQGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
 QY 121 LVFGLTANSPTHLIQGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
 DB 121 LVFGLTANSPTHLIQGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDG 180
 QY 181 TWCTVLONOKKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240
 DB 181 TWCTVLONOKKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240
 QY 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLQMGKKLPLHLTLPQALPOYAGSGNLTIA 300
 DB 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLQMGKKLPLHLTLPQALPOYAGSGNLTIA 300
 QY 301 LEATGTGLHGEVNLVVMRATQLOKNTLCEWGPSPKLMSTLKENKEAVSKREKPVWV 360
 DB 301 LEATGTGLHGEVNLVVMRATQLOKNTLCEWGPSPKLMSTLKENKEAVSKREKPVWV 360
 QY 361 LNPAGMOCCLSDSGOVLLESNIKVLPTWSTVPAPAPKSC 404
 DB 361 LNPAGMOCCLSDSGOVLLESNIKVLPTWSTVPAPAPKSC 404

```
RESULT 27
US-08-284-391B-5
; Sequence 5, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391B
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-284-391B-5

Query Match      59.5%; Score 2032.5; DB 2; Length 462;
Best Local Similarity 98.0%; Pred. No. 2.6e-153;
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;
```

```

QY 241 QAEKASSKSWITFDLKNKEVSVKRVTDPPKLOMGKLLPLHLTPQALPOYAGSGNLTIA 300
DB 241 QAEKASSKSWITFDLKNKEVSVKRVTDPPKLOMGKLLPLHLTPQALPOYAGSGNLTIA 300
QY 301 LEAKTGKLEHVEVNLVVMRATLOKULTCFVWGPTSPKMLSLKLENKAKVSRKRPVWY 360
DB 301 LEAKTGKLEHVEVNLVVMRATLOKULTCFVWGPTSPKMLSLKLENKAKVSRKRPVWY 360
QY 361 LNPBAGMOCLLSDSGQVLLSEINIKVLPWTWSTPVPAPAPKSC 404
DB 361 LNPBAGMOCLLSDSGQVLLSEINIKVLPWTWSTPVAH--DPOLC 401

RESULT 28
US-09-218-950-5
; Sequence 5, Application US/09218950
; Patent No. 6284240
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/218,950
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-218-950-5

Query Match      59.5%; Score 2032.5; DB 3; Length 462;
Best Local Similarity 98.0%; Pred. No. 2.6e-153;
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;
```

Db 1 MNRGVPFRHLLVLVQLALLPAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNOIK 60
Qy 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
Db 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
Qy 121 LVFGLTANSDFHLLQGGSLTLTLSPGSSPSVOCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSDFHLLQGGSLTLTLSPGSSPSVOCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy 181 TWCTVLONQKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPFLATFVEKLTGSGELMW 240
Db 181 TWCTVLONQKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPFLATFVEKLTGSGELMW 240
Qy 241 QAERASSKSWITFDLKNKEVSVKRVTPDPLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
Db 241 QAERASSKSWITFDLKNKEVSVKRVTPDPLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
Qy 301 LEATGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKKEKPYMV 360
Db 301 LEATGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKKEKPYMV 360
Qy 361 LNPAAGMWQCLLSDSGVLLSNIKVLPTWSTPVPCEAPBPKSC 404
Db 361 LNPAAGMWQCLLSDSGVLLSNIKVLPTWSTPVPCEAPBPKSC 404

RESULT 29
PCT-US92-01785-5
Sequence 5, Application PC/TUS9201785

GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" diskette, 1.44 Mb
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-01785-5

Query Match 59.5%; Score 2032.5; DB 5; Length 462;
Best Local Similarity 98.0%; Pred. No. 2,6e-153;
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

Qy 1 MNRGVPFRHLLVLVQLALLPAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNOIK 60
Db 1 MNRGVPFRHLLVLVQLALLPAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNNOIK 60
Qy 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
Db 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
Qy 121 LVFGLTANSDFHLLQGGSLTLTLSPGSSPSVOCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSDFHLLQGGSLTLTLSPGSSPSVOCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy 181 TWCTVLONQKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPFLATFVEKLTGSGELMW 240
Db 181 TWCTVLONQKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPFLATFVEKLTGSGELMW 240
Qy 241 QAERASSKSWITFDLKNKEVSVKRVTPDPLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
Db 241 QAERASSKSWITFDLKNKEVSVKRVTPDPLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
Qy 301 LEATGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKKEKPYMV 360
Db 301 LEATGKLGHEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKKEKPYMV 360
Qy 361 LNPAAGMWQCLLSDSGVLLSNIKVLPTWSTPVPCEAPBPKSC 404
Db 361 LNPAAGMWQCLLSDSGVLLSNIKVLPTWSTPVPCEAPBPKSC 404

RESULT 30

PCT-US95-00454-5
Sequence 5, Application PC/TUS9500454

GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00454-5

Query Match 59.5%; Score 2032.5; DB 5; Length 462;
Best Local Similarity 98.0%; Pred. No. 2.5e-153;
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVFRRLLVLTQALPAATQGNKVVLGKKGDVETLTCTASQKKSIOFHMKNSNQIK 60
DB 1 MNRGVFRRLLVLTQALPAATQGNKVVLGKKGDVETLTCTASQKKSIOFHMKNSNQIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIIKNKIEDSDTYICEVEDQKEEVL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIIKNKIEDSDTYICEVEDQKEEVL 120
QY 121 LVFGLTANSDTHLLOQGSLLTLTLESPGSSPSVQCSPRGKNIQGGKTLSVSQLLEQDSG 180
DB 121 LVFGLTANSDTHLLOQGSLLTLTLESPGSSPSVQCSPRGKNIQGGKTLSVSQLLEQDSG 180
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSFPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSFPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKKLPLHLTLPQALPOYAGSGNLTLA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKKLPLHLTLPQALPOYAGSGNLTLA 300
QY 301 LEAKTGKLEHENVLVVWRATOLQKNLTCEWGPSTPKMLSLKLENKEAKVSKREKPVWV 360
DB 301 LEAKTGKLEHENVLVVWRATOLQKNLTCEWGPSTPKMLSLKLENKEAKVSKREKPVWV 360
QY 361 LNPEAGMWQCLSDSGQVLLSENIKVLPTWSTPVPCPAPEPKSC 404
DB 361 LNPEAGMWQCLSDSGQVLLSENIKVLPTWSTPVHA---DPOIC 401

RESULT 31
US-08-284-391B-29
Sequence 29, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-29

Query Match 59.4%; Score 2029; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 4e-153;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVFRRLLVLTQALPAATQGNKVVLGKKGDVETLTCTASQKKSIOFHMKNSNQIK 60
DB 1 MNRGVFRRLLVLTQALPAATQGNKVVLGKKGDVETLTCTASQKKSIOFHMKNSNQIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIIKNKIEDSDTYICEVEDQKEEVL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIIKNKIEDSDTYICEVEDQKEEVL 120
QY 121 LVFGLTANSDTHLLOQGSLLTLTLESPGSSPSVQCSPRGKNIQGGKTLSVSQLLEQDSG 180
DB 121 LVFGLTANSDTHLLOQGSLLTLTLESPGSSPSVQCSPRGKNIQGGKTLSVSQLLEQDSG 180
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSFPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSFPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKKLPLHLTLPQALPOYAGSGNLTLA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKKLPLHLTLPQALPOYAGSGNLTLA 300
QY 301 LEAKTGKLEHENVLVVWRATOLQKNLTCEWGPSTPKMLSLKLENKEAKVSKREKPVWV 360
DB 301 LEAKTGKLEHENVLVVWRATOLQKNLTCEWGPSTPKMLSLKLENKEAKVSKREKPVWV 360
QY 361 LNPEAGMWQCLSDSGQVLLSENIKVLPTWSTPV 394
DB 361 LNPEAGMWQCLSDSGQVLLSENIKVLPTWSTPV 394

RESULT 32
US-09-218-950-29
Sequence 29, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/218,950
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-218-950-29

```

```

Query Match      59.4%; Score 2029; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 4e-153;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPRHLLVQLALPPAATQGNKVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
DB 1 MNRGVPRHLLVQLALPPAATQGNKVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLOQGSLLTLTSPGSSPSVQCRRPRGKNTIOGKTLSSVQLELDDSG 180
DB 121 LVFGLTANSDTHLLOQGSLLTLTSPGSSPSVQCRRPRGKNTIOGKTLSSVQLELDDSG 180
QY 181 TWCTCTVQONOKVEFKIDIVLAFQKASSIYKKKEGOVERSPFLAFTVEKLTGSGELMW 240
DB 181 TWCTCTVQONOKVEFKIDIVLAFQKASSIYKKKEGOVERSPFLAFTVEKLTGSGELMW 240
QY 241 QAEARASSSKSWITFDLKNKEVSVKRVTQDPKLGKGLPLHLTLPOALPOYAGSGNLTLLA 300
DB 241 QAEARASSSKSWITFDLKNKEVSVKRVTQDPKLGKGLPLHLTLPOALPOYAGSGNLTLLA 300
QY 301 LEAKTGLHGEVNLVVMRATQLOKNTLCEVWGPTSPKLMSTLKENKEAVSKREKPVVW 360
DB 301 LEAKTGLHGEVNLVVMRATQLOKNTLCEVWGPTSPKLMSTLKENKEAVSKREKPVVW 360
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPV 394
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPV 394

```

```

RESULT 33
US-08-466-368-4
; Sequence 4, Application US/08466368
; Patent No. 6083539
; GENERAL INFORMATION:
; APPLICANT: Maddon, Paul J.
; APPLICANT: Littman, Dan R.
; APPLICANT: Cheese, Leonard

```

```

; APPLICANT: Axel, Richard
; APPLICANT: Weiss, Robin
; APPLICANT: McDougall, J. S.
; TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,368
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 24577-EI-B/JPM/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 458 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-368-4

```

```

Query Match      59.4%; Score 2029; DB 3; Length 458;
Best Local Similarity 100.0%; Pred. No. 4.9e-153;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPRHLLVQLALPPAATQGNKVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
DB 1 MNRGVPRHLLVQLALPPAATQGNKVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLOQGSLLTLTSPGSSPSVQCRRPRGKNTIOGKTLSSVQLELDDSG 180
DB 121 LVFGLTANSDTHLLOQGSLLTLTSPGSSPSVQCRRPRGKNTIOGKTLSSVQLELDDSG 180
QY 181 TWCTCTVQONOKVEFKIDIVLAFQKASSIYKKKEGOVERSPFLAFTVEKLTGSGELMW 240
DB 181 TWCTCTVQONOKVEFKIDIVLAFQKASSIYKKKEGOVERSPFLAFTVEKLTGSGELMW 240
QY 241 QAEARASSSKSWITFDLKNKEVSVKRVTQDPKLGKGLPLHLTLPOALPOYAGSGNLTLLA 300
DB 241 QAEARASSSKSWITFDLKNKEVSVKRVTQDPKLGKGLPLHLTLPOALPOYAGSGNLTLLA 300
QY 301 LEAKTGLHGEVNLVVMRATQLOKNTLCEVWGPTSPKLMSTLKENKEAVSKREKPVVW 360
DB 301 LEAKTGLHGEVNLVVMRATQLOKNTLCEVWGPTSPKLMSTLKENKEAVSKREKPVVW 360
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPV 394
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPV 394

```

```

RESULT 34
US-08-328-500-9
; Sequence 9, Application US/08328500

```

Patent No. 6673896
GENERAL INFORMATION:
APPLICANT: Maddon, Paul J.
APPLICANT: Axel, Richard
APPLICANT: Sweet, Richard W.
APPLICANT: Archos, James
TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/328,500
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/24577-CY
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ. ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 457 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-328-500-9

Query Match 59.2%; Score 2021; DB 4; Length 457;
Best Local Similarity 99.7%; Pred. No. 2.1e-152;

Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTALLPATQGNKVVLGKKGDTVELTCTASOKKSIOFHMKSNQK 60
DB 1 MNRGVPRHLLVLTALLPATQGNKVVLGKKGDTVELTCTASOKKSIOFHMKSNQK 60
QY 61 ILGNQSFLLTKGSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVL 120
DB 61 ILGNQSFLLTKGSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVL 120
QY 121 LVFGLTANSPTLHLLQGOSLTLLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPTLHLLQGOSLTLLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKGEGVEFPLAFYVEKLTGSGELMW 240
DB 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKGEGVEFPLAFYVEKLTGSGELMW 240
QY 241 QARRASSSKSWITFDLKNKEVSVKRVYQDPKLGKGLPLHLTLPOLPOYAGSGNLTLLA 300
DB 241 QARRASSSKSWITFDLKNKEVSVKRVYQDPKLGKGLPLHLTLPOLPOYAGSGNLTLLA 300
QY 301 LEAKTGKLGHOEVNLVVMRATQLOKNLTCEVWGSPSPKMLSLKLENKAVSRREKPVWV 360
DB 301 LEAKTGKLGHOEVNLVVMRATQLOKNLTCEVWGSPSPKMLSLKLENKAVSRREKPVWV 360
QY 361 LNPBAGMOCLLSDSGVLLSNIKVLPPTSTPV 394
DB 361 LNPBAGMOCLLSDSGVLLSNIKVLPPTSTPV 394

RESULT 35

US-08-236-311-1
Sequence 1, Application US/08236311
Patent No. 5565335

GENERAL INFORMATION:
APPLICANT: Capon, Daniel J.
APPLICANT: Gregory, Timothy J.
TITLE OF INVENTION: Adhesion Variants
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/236,311
FILING DATE: 02-MAY-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936190
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842777
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/250785
FILING DATE: 28-SEP-1988

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/104329
FILING DATE: 02-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 444P1C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881

TELEX: 910/371-7168
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-236-311-1

Query Match 59.1%; Score 2017; DB 1; Length 402;
Best Local Similarity 99.7%; Pred. No. 3.6e-152;

Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTALLPATQGNKVVLGKKGDTVELTCTASOKKSIOFHMKSNQK 60
DB 1 MNRGVPRHLLVLTALLPATQGNKVVLGKKGDTVELTCTASOKKSIOFHMKSNQK 60
QY 61 ILGNQSFLLTKGSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVL 120
DB 61 ILGNQSFLLTKGSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVL 120
QY 121 LVFGLTANSPTLHLLQGOSLTLLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPTLHLLQGOSLTLLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKGEGVEFPLAFYVEKLTGSGELMW 240
DB 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKGEGVEFPLAFYVEKLTGSGELMW 240
QY 241 QARRASSSKSWITFDLKNKEVSVKRVYQDPKLGKGLPLHLTLPOLPOYAGSGNLTLLA 300
DB 241 QARRASSSKSWITFDLKNKEVSVKRVYQDPKLGKGLPLHLTLPOLPOYAGSGNLTLLA 300

Db 241 QABASSSKSMITFDLKNKEVSVRKVTDPKLOMGKKLPLHLTLPALPOYAGSGNLTLA 300
Qy 301 LEATGKLGHOEVLVVRATOLQKNTCEVWGPTSPKLMSTLKENKEAKVSKREKPYWV 360
Db 301 LEATGKLGHOEVLVVRATOLQKNTCEVWGPTSPKLMSTLKENKEAKVSKREKAVWV 360
Qy 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTP 393
Db 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTP 393

RESULT 36
US-08-457-918-1
Sequence 1, Application US/08457918
Patent No. 6117655
GENERAL INFORMATION:
APPLICANT: Capon, Daniel J.
APPLICANT: Gregory, Timothy J.
TITLE OF INVENTION: Adhesion Variance
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 Inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,918
FILING DATE: 1-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/236311
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936190
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842777
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/250785
FILING DATE: 28-SEP-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/104329
FILING DATE: 02-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Kubinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0444P1C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8228
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-457-918-1

Query March 59.1%; Score 2017; DB 3; Length 402;
Best Local Similarity 99.7%; Pred. No. 3.6e-152;
Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MNRGVPRHLLLVQLALLPATQGNKVVLOKGDVVELTCTASOKSIOFHMKNNOIK 60
Db 1 MNRGVPRHLLLVQLALLPATQGNKVVLOKGDVVELTCTASOKSIOFHMKNNOIK 60

Qy 61 ILGNQGSFLTRKGPSKLNDRADSRSLWMDQNFPIIKNLKIEDSDTYICEVEDOKEEVOL 120
Db 61 ILGNQGSFLTRKGPSKLNDRADSRSLWMDQNFPIIKNLKIEDSDTYICEVEDOKEEVOL 120
Qy 121 LVFGLTANSDTHLLQGGSLTLLTSPGSSPSVQCRSPRKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSDTHLLQGGSLTLLTSPGSSPSVQCRSPRKNIQGGKTLVSQLELDQSG 180
Qy 181 TWCTTVQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLATVEKLTGSGELMW 240
Db 181 TWCTTVQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLATVEKLTGSGELMW 240
Qy 241 QABASSSKSMITFDLKNKEVSVRKVTDPKLOMGKKLPLHLTLPALPOYAGSGNLTLA 300
Db 241 QABASSSKSMITFDLKNKEVSVRKVTDPKLOMGKKLPLHLTLPALPOYAGSGNLTLA 300
Qy 301 LEATGKLGHOEVLVVRATOLQKNTCEVWGPTSPKLMSTLKENKEAKVSKREKPYWV 360
Db 301 LEATGKLGHOEVLVVRATOLQKNTCEVWGPTSPKLMSTLKENKEAKVSKREKAVWV 360
Qy 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTP 393
Db 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTP 393

RESULT 37
US-09-039-555B-15
Sequence 15, Application US/09039555B
Patent No. 6033856
GENERAL INFORMATION:
APPLICANT: Koerner, Kathrin
APPLICANT: Mueller, Rolf
APPLICANT: Sadiacek, Hans-Harald
TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,555B
FILING DATE: 16-MAR-1998
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 19710643.9
FILING DATE: 14-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 016779/0131
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 458 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-039-555B-15

Query Match 59.0%; Score 2015; DB 3; Length 458;
 Best Local Similarity 99.5%; Pred. No. 6,3e-152;
 Matches 392; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNRGVFRHLVLVQLALLPAATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60
 DB 1 MNRGVFRHLVLVQLALLPAATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60
 QY 61 ILGNQSSFLTKGPKSKNDRAADRSRSLMDQGNFPLIIRNKLIEDSDTYICEVEDQKEEVL 120
 DB 61 ILGNQSSFLTKGPKSKNDRAADRSRSLMDQGNFPLIIRNKLIEDSDTYICEVEDQKEEVL 120
 QY 121 LVFGLTANSDBTHLLOQOSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 DB 121 LVFGLTANSDBTHLLOQOSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 QY 181 TWCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
 DB 181 TWCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
 QY 241 QAERASSSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPQALPOYAGSGLTLA 300
 DB 241 QAERASSSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPQALPOYAGSGLTLA 300
 QY 301 LEAKTKLHOEVNLVWMBATOLQKNLTCEVWGPTSPKMLSLKLNKEAKVSKREKPYWV 360
 DB 301 LEAKTKLHOEVNLVWMBATOLQKNLTCEVWGPTSPKMLSLKLNKEAKVSKREKPYWV 360
 QY 361 LNPEAGMWQCLISDSGOVLLESNIKVLPTWSTPV 394
 DB 361 LNPEAGMWQCLISDSGOVLLESNIKVLPTWSTPV 394

RESULT 38
 US-09-517-605-3
 ; Sequence 3, Application US/09517605
 ; Patent No. 6391567
 ; GENERAL INFORMATION:
 ; APPLICANT: Littman, Dan R.
 ; APPLICANT: Kwon, Douglas S.
 ; APPLICANT: van Kooyk, Yvette
 ; APPLICANT: Geijtenbeek, Theo
 ; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
 ; FILE OF INVENTION: CELLS
 ; FILE REFERENCE: 1049-1-017
 ; CURRENT APPLICATION NUMBER: US/09/517,605
 ; CURRENT FILING DATE: 2000-03-02
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 458
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-517-605-3

Query Match 58.8%; Score 2007; DB 4; Length 458;
 Best Local Similarity 99.2%; Pred. No. 2,7e-151;
 Matches 391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNRGVFRHLVLVQLALLPAATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60
 DB 1 MNRGVFRHLVLVQLALLPAATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60
 QY 61 ILGNQSSFLTKGPKSKNDRAADRSRSLMDQGNFPLIIRNKLIEDSDTYICEVEDQKEEVL 120
 DB 61 ILGNQSSFLTKGPKSKNDRAADRSRSLMDQGNFPLIIRNKLIEDSDTYICEVEDQKEEVL 120
 QY 121 LVFGLTANSDBTHLLOQOSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 DB 121 LVFGLTANSDBTHLLOQOSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 QY 181 TWCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
 DB 181 TWCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240

DB 181 TWCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
 QY 241 QAERASSSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPQALPOYAGSGLTLA 300
 DB 241 QAERASSSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPQALPOYAGSGLTLA 300
 QY 301 LEAKTKLHOEVNLVWMBATOLQKNLTCEVWGPTSPKMLSLKLNKEAKVSKREKPYWV 360
 DB 301 LEAKTKLHOEVNLVWMBATOLQKNLTCEVWGPTSPKMLSLKLNKEAKVSKREKPYWV 360
 QY 361 LNPEAGMWQCLISDSGOVLLESNIKVLPTWSTPV 394
 DB 361 LNPEAGMWQCLISDSGOVLLESNIKVLPTWSTPV 394

RESULT 39
 US-08-466-368-2
 ; Sequence 2, Application US/08466368
 ; Patent No. 6093539
 ; GENERAL INFORMATION:
 ; APPLICANT: Madden, Paul J.
 ; APPLICANT: Littman, Dan R.
 ; APPLICANT: Chess, Leonard
 ; APPLICANT: Axel, Richard
 ; APPLICANT: Weiss, Robin
 ; APPLICANT: McDougall, J. S.
 ; TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
 ; TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cooper & Dunham LLP
 ; STREET: 1185 Avenue of Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/466,368
 ; FILING DATE: 06-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P.
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 24577-EI-B/JPW/AKC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-278-0400
 ; TELEFAX: 212-391-0525
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 394 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-466-368-2

Query Match 58.6%; Score 2001; DB 3; Length 394;
 Best Local Similarity 98.7%; Pred. No. 6,6e-151;
 Matches 389; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNRGVFRHLVLVQLALLPAATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60
 DB 1 MNRGVFRHLVLVQLALLPAATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60
 QY 61 ILGNQSSFLTKGPKSKNDRAADRSRSLMDQGNFPLIIRNKLIEDSDTYICEVEDQKEEVL 120
 DB 61 ILGNQSSFLTKGPKSKNDRAADRSRSLMDQGNFPLIIRNKLIEDSDTYICEVEDQKEEVL 120
 QY 121 LVFGLTANSDBTHLLOQOSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 DB 121 LVFGLTANSDBTHLLOQOSLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180

Db 121 LVFGLTANSDTHLQOGSLTTLTLESPGSSSPVQCRSPRGKNIQGGKTLTSVSOLELQDSG 180
|||
Qy 181 TWCTVLONOKKVEFKIDIVVLAFOKASSIYKKEGEVDFSPFLAFTVEKLTSGGELMW 240
|||
Db 181 TWCTVLONOKKVEFKIDIVVLAFOKASSIYKKEGEVDFSPFLAFTVEKLTSGGELMW 240
|||
Qy 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
|||
Db 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
|||
Qy 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMLSLKLENKAKVSKREKPVW 360
|||
Db 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMLSLKLENKAKVSKREKPVW 360
|||
Qy 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394
|||
Db 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394
|||

RESULT 40

US-08-328-500-2
; Sequence 2, Application US/08328500
; Patent No. 6673896
; GENERAL INFORMATION:
; APPLICANT: Maddon, Paul J.
; APPLICANT: Axel, Richard
; APPLICANT: Sweet, Richard W.
; APPLICANT: Athos, James
; TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,500
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/24577-CY
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-328-500-2

Query Match 58.6%; Score 2001; DB 4; Length 394;
Best Local Similarity 98.7%; Pred. No. 6.6e-151;

Matches 389; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MNRGVPRHLLLVQLALLPATQGNKYVLGKGGDTVELTCTASOKSIOFHMKNSNOK 60
|||
Db 1 MNRGVPRHLLLVQLALLPATQGNKYVLGKGGDTVELTCTASOKSIOFHMKNSNOK 60
|||
Qy 61 ILGNQGSFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
|||
Db 61 ILGNQGSFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
|||

Qy 121 LVFGLTANSDTHLQOGSLTTLTLESPGSSSPVQCRSPRGKNIQGGKTLTSVSOLELQDSG 180
|||
Db 121 LVFGLTANSDTHLQOGSLTTLTLESPGSSSPVQCRSPRGKNIQGGKTLTSVSOLELQDSG 180
|||
Qy 181 TWCTVLONOKKVEFKIDIVVLAFOKASSIYKKEGEVDFSPFLAFTVEKLTSGGELMW 240
|||
Db 181 TWCTVLONOKKVEFKIDIVVLAFOKASSIYKKEGEVDFSPFLAFTVEKLTSGGELMW 240
|||
Qy 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
|||
Db 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
|||
Qy 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMLSLKLENKAKVSKREKPVW 360
|||
Db 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMLSLKLENKAKVSKREKPVW 360
|||
Qy 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394
|||
Db 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394
|||

RESULT 41

5223394-7
; Patent No. 5223394
; APPLICANT: WALLNER, BARBARA
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
; LINKAGE SIGNAL SEQUENCE
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/335,688
; FILING DATE: 10-APR-1989
; SEQ ID NO: 7
; LENGTH: 458
5223394-7

Query Match 58.3%; Score 1989; DB 6; Length 458;
Best Local Similarity 98.5%; Pred. No. 7.4e-150;

Matches 388; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MNRGVPRHLLLVQLALLPATQGNKYVLGKGGDTVELTCTASOKSIOFHMKNSNOK 60
|||
Db 1 MNRGVPRHLLLVQLALLPATQGNKYVLGKGGDTVELTCTASOKSIOFHMKNSNOK 60
|||
Qy 61 ILGNQGSFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
|||
Db 61 ILGNQGSFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
|||
Qy 121 LVFGLTANSDTHLQOGSLTTLTLESPGSSSPVQCRSPRGKNIQGGKTLTSVSOLELQDSG 180
|||
Db 121 LVFGLTANSDTHLQOGSLTTLTLESPGSSSPVQCRSPRGKNIQGGKTLTSVSOLELQDSG 180
|||
Qy 181 TWCTVLONOKKVEFKIDIVVLAFOKASSIYKKEGEVDFSPFLAFTVEKLTSGGELMW 240
|||
Db 181 TWCTVLONOKKVEFKIDIVVLAFOKASSIYKKEGEVDFSPFLAFTVEKLTSGGELMW 240
|||
Qy 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
|||
Db 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
|||
Qy 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMLSLKLENKAKVSKREKPVW 360
|||
Db 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMLSLKLENKAKVSKREKPVW 360
|||
Qy 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394
|||
Db 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394
|||

RESULT 42
5223418-2
; Patent No. 5223418

```

; APPLICANT: ARCURI, EDWARD J.;BRAUNER, MARY E.;DONOVAN, MARY
; J.; GERBER, ROBERT G.; KELLER, JOHN A.
; TITLE OF INVENTION: METHOD OF IMPROVING THE YIELD OF
; HETEROLOGOUS PROTEINS PRODUCED BY STREPTOMYCES LIVIDANS
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/589,979
; FILING DATE: 28-SEP-1990
; SEQ ID NO: 2
; LENGTH: 394
5223418-2

Query Match      57.1%; Score 1951; DB 6; Length 394;
Best Local Similarity 97.5%; Pred. No. 6.2e-147;
Matches 384; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      1 MNRGVPFRHLVLVQLALPRAATQGNKVVGLKGGDTVELTCTASQKKSIOFHMKNNOIK 60
DB      1 MNRGVPFRHLVLVQLALPRAATQGNKVVGLKGGDTVELTCTASQKKSIOFHMKNNOIK 60
QY      61 ILGNQGSFLTGGPSKLNDRADSRSLMDQGNFPLITKNLKIEDSDTYICEVEDQKEBVOL 120
DB      61 ILGNQGSFLTGGPSKLNDRADSRSLMDQGNFPLITKNLKIEDSDTYICEVEDQKEBVOL 120
QY      121 LVFGLTANSPDTHLLOGSLTLTLESPPGSSPVQCSPPRGKNIQGGKTLISVSOLELQDSG 180
DB      121 LVFGLTANSPDTHLLOGSLTLTLESPPGSSPVQCSPPRGKNIQGGKTLISVSOLELQDSG 180
QY      181 TWTCTVLQNGKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240
DB      181 TWTCTVLQNGKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240
QY      241 QABRASSKSWITFDLKNKEVSVKRVTDQPKLQMGKPLHLTLPOALPOYAGSGNLTTLA 300
DB      241 QABRASSKSWITFDLKNKEVSVKRVTDQPKLQMGKPLHLTLPOALPOYAGSGNLTTLA 300
QY      301 LEAKTGKLEHGVNLVVMRATQLOKNLCEVWGPTSPKLMSTLKENKEAKVSKREKPVWV 360
DB      301 LEAKTGKLEHGVNLVVMRATQLOKNLCEVWGPTSPKLMSTLKENKEAKVSKREKPVWV 360
QY      361 LNPBAGMOCCLSDSGQVLLSENIKVLPTWSTP 394
DB      361 LNPBAGMOCCLSDSGQVLLSENIKVLPTWSTP 394

RESULT 43
US-08-236-311-4
; Sequence 4, Application US/08236311
; Patent No. 5565335
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,311
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 444P1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-236-311-4

Query Match      55.8%; Score 1904; DB 1; Length 434;
Best Local Similarity 99.7%; Pred. No. 3.9e-143;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      24 QGNKVVGLKGGDTVELTCTASQKKSIOFHMKNNOIKILGNQGSFLTGGPSKLNDRADR 83
DB      56 QGNKVVGLKGGDTVELTCTASQKKSIOFHMKNNOIKILGNQGSFLTGGPSKLNDRADR 115
QY      84 RSLMDQGNFPLITKNLKIEDSDTYICEVEDQKEBVOLLVFGLTANSPDTHLLOGSLTLTL 143
DB      116 RSLMDQGNFPLITKNLKIEDSDTYICEVEDQKEBVOLLVFGLTANSPDTHLLOGSLTLTL 175
QY      144 ESPPGSSPVQCSPPRGKNIQGGKTLISVSOLELQDSGTLTCTVLQNGKVEFKIDIVLVA 203
DB      176 ESPPGSSPVQCSPPRGKNIQGGKTLISVSOLELQDSGTLTCTVLQNGKVEFKIDIVLVA 235
QY      204 FQKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMWQABRASSKSWITFDLKNKEVSV 263
DB      236 FQKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMWQABRASSKSWITFDLKNKEVSV 295
QY      264 KRVTDQPKLQMGKPLHLTLPOALPOYAGSGNLTALAKTKGLHGVNLVVMRATQLO 323
DB      296 KRVTDQPKLQMGKPLHLTLPOALPOYAGSGNLTALAKTKGLHGVNLVVMRATQLO 355
QY      324 KNLTCVWGPTSPKLMSTLKENKEAKVSKREKPVWVNLNPBAGMOCCLSDSGQVLLSENI 383
DB      356 KNLTCVWGPTSPKLMSTLKENKEAKVSKREKPVWVNLNPBAGMOCCLSDSGQVLLSENI 415
QY      384 IKVLPTWSTP 393
DB      416 IKVLPTWSTP 425

RESULT 44
US-08-457-918-4
; Sequence 4, Application US/08457918
; Patent No. 6117655
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,918
FILING DATE: 1-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/236311
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936190
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842777
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/250785
FILING DATE: 28-SEP-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/104329
FILING DATE: 02-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Kubinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0444P1C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8228
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 434 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-457-918-4

```

QY 386 VLPTWSTPV 394
| | | | |
DB 361 VLPTWSTPV 369

RESULT 46

US-08-808-374-1
; Sequence 1, Application US/08808374
; Patent No. 5961976
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang YI
; TITLE OF INVENTION: Antibody Against a Host Cell
; TITLE OF INVENTION: Antigen Complex for Pre- and Post-Exposure
; TITLE OF INVENTION: Protection from Infection by HIV Primary Isolates
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,374
; FILING DATE: 28-Feb-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/657,149
; FILING DATE: 03-June-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-808-374-1

Query Match 55.4%; Score 1892; DB 2; Length 433;
Best Local Similarity 99.5%; Pred. No. 3, Se-142;
Matches 367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 26 NKVLGKGGTVELTCTASQKKSIOFWKNSNOIKIINGOSFLTTPSKLNDADSRRL 85
| | | | |
DB 1 NKVLGKGGTVELTCTASQKKSIOFWKNSNOIKIINGOSFLTTPSKLNDADSRRL 60
| | | | |
QY 86 LMOGNSPLIIKLIKIDSDTYICEVEDQKEVOLVFGLTANSDDLLOGOSLTTLLES 145
| | | | |
DB 61 LMOGNSPLIIKLIKIDSDTYICEVEDQKEVOLVFGLTANSDDLLOGOSLTTLLES 120
| | | | |
QY 146 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDOSGTWCTVLONQKVEFKIDIVIAFQ 205
| | | | |
DB 121 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDOSGTWCTVLONQKVEFKIDIVIAFQ 180
| | | | |
QY 206 KASSIVYKKEGEVSEFPLAFYVEKLTGSGELMWOABRASSSSSWITFDLKNEVSVKR 255
| | | | |
DB 181 KASSIVYKKEGEVSEFPLAFYVEKLTGSGELMWOABRASSSSSWITFDLKNEVSVKR 240
| | | | |
QY 266 VTGDPKLQMGKTLPLHLTLPOALPOVAGSGNLTALFAKTKLHOEVNLVVMRATOLQKN 325
| | | | |

DB 241 VTGDPKLQMGKTLPLHLTLPOALPOVAGSGNLTALFAKTKLHOEVNLVVMRATOLQKN 300
| | | | |
QY 326 LTCCEWGGPTSPKLMSTIKLENKAKSKREKVVWVNLNPAQWQCLSSGOVLTLESNIK 385
| | | | |
DB 301 LTCCEWGGPTSPKLMSTIKLENKAKSKREKVVWVNLNPAQWQCLSSGOVLTLESNIK 360
| | | | |
QY 386 VLPTWSTPV 394
| | | | |
DB 361 VLPTWSTPV 369

RESULT 47

US-09-100-409A-1
; Sequence 1, Application US/09100409A
; Patent No. 6030388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang YI
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-1

Query Match 55.3%; Score 1887; DB 3; Length 433;
Best Local Similarity 99.5%; Pred. No. 8, 7e-142;
Matches 366; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 27 KVLGKGGTVELTCTASQKKSIOFWKNSNOIKIINGOSFLTTPSKLNDADSRRL 86
| | | | |
DB 2 KVLGKGGTVELTCTASQKKSIOFWKNSNOIKIINGOSFLTTPSKLNDADSRRL 61
| | | | |
QY 87 WDOGNSPLIIKLIKIDSDTYICEVEDQKEVOLVFGLTANSDDLLOGOSLTTLLES 146
| | | | |
DB 62 WDOGNSPLIIKLIKIDSDTYICEVEDQKEVOLVFGLTANSDDLLOGOSLTTLLES 121
| | | | |
QY 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDOSGTWCTVLONQKVEFKIDIVIAFQ 206
| | | | |
DB 122 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDOSGTWCTVLONQKVEFKIDIVIAFQ 181
| | | | |
QY 207 ASSIVYKKEGEVSEFPLAFYVEKLTGSGELMWOABRASSSSSWITFDLKNEVSVKR 266
| | | | |
DB 182 ASSIVYKKEGEVSEFPLAFYVEKLTGSGELMWOABRASSSSSWITFDLKNEVSVKR 241
| | | | |


```

; Sequence 17, Application US/09375419
; Patent No. 6264950
; GENERAL INFORMATION:
; APPLICANT: Staerz, Uwe
; TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
; TITLE OF INVENTION: LYMPHOCYTE VERO
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, 35th Floor
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/375,419
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,172
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2879-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-375-419-17

Query Match      46.4%; Score 1585.5; DB 3; Length 410;
Best Local Similarity 54.2%; Pred. No. 7,3e-118;
Matches 32; Conservative 33; Mismatches 51; Indels 191; Gaps 6;

QY      26 NKVVLGGKGDVTELTCTASQKSIQFHWKNSNOIKITGNQGSFLTKGSPSKLNDRADSRRS 85
DB      1 NKVVLGGKGDVTELTCTASQKSIQFHWKNSNOIKITGNQGSFLTKGSPSKLNDRADSRRS 60

QY      86 LMOGNGPRLTIKULKIEDSDTYICEVEDQKEVQLVFGLTANSPTHLLQGOSLTLTLES 145
DB      61 LMOGNGPRLTIKULKIEDSDTYICEVEDQKEVQLVFGLTANSPTHLLQGOSLTLTLES 120

QY      146 PGGSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLONQKVEFKIDIVLAFQ 205
DB      121 PGGSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLONQKVEFKIDIVLAFQ 178

QY      206 KASSIVYKKEGEVQSFPLAFVTEKLTGSGELMWMQERASSSKSWITFDLKNKEVSVKR 265
DB      179 ----- 178

QY      266 VTQDPKIQMGKULPLHLTLPOALPOYAGSGNLTALAEAKTGKLAHQEVNLVWMRAIOLQKN 325
DB      179 ----- 181

QY      326 LTCEVWGPSTPKMLSLKLENKAKVSKREKPVVVLNPEAGMQLCLSDSGQVLTLESNIK 385
DB      182 -----IK 186

QY      386 VLEPTWSTVPCPAPBPKSCDKTHTCPCLLGGPSVYLFPPPKDTLMISRTPEVTCVVVDV 445
DB      187 PCP-----PCKCPA-----PMLGGPSVYLFPPPKDKVLMISLSPITVTCVVVDV 230

```

```

QY      446 SHEDEPKENMYVDGVEVHNAKTKPREBOYNSTYRVSVLTVLHODMLNGKEYCKCVSNK 505
DB      231 SEDDEPVOISWFEVNVNVEHTAQTQREDYNSRLKAVSLPQHQDMMGKCKKCVNNK 290

QY      506 ALPAPIEKTISSKAKOPREPOVYTLPPSRDELTKNOVSLTCKVGFPSDIAVAVESNQ 565
DB      291 DLPAPIERTISKPKSVRAFYVYVLPPE-EEWTKQVTLTCVTDPMFEDIYVEWTNNK 349

QY      566 PENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGVNFSVNHAEAHNYTQKSLSLSPG 625
DB      350 TELNYKNTPEVLDSDGSFYFMSKLVKKNWERNSSCSVHHEGLHNHTTKSFSTRG 409

RESULT 51
5223394-11
; Patent No. 5223394
; APPLICANT: WALLNER, BARBARA
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
; LINKAGE SIGNAL SEQUENCE
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/335,688
; FILING DATE: 10-APR-1989
; SEQ ID NO:11:
; LENGTH: 318

Query Match      40.1%; Score 1368; DB 6; Length 318;
Best Local Similarity 93.4%; Pred. No. 9,7e-101;
Matches 268; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY      1 NMRGVFRRLLLVLTQALLPAATQGNKVTLGKKGDTVELTCTASQKSIQFHWKNSNOIK 60
DB      1 NMRGVFRRLLLVLTQALLPAATQGNKVTLGKKGDTVELTCTASQKSIQFHWKNSNOIK 60

QY      61 IIGNQSFLTKGSPSKLNDRADSRRLMDQGNFPLTIKULKIEDSDTYICEVEDQKEVQL 120
DB      61 IIGNQSFLTKGSPSKLNDRADSRRLMDQGNFPLTIKULKIEDSDTYICEVEDQKEVQL 120

QY      121 LVFGLTANSPTHLLQGOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB      121 LVFGLTANSPTHLLQGOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY      181 TWTCTVLONQKVEFKIDIVLAFQKASSIVYKKEGEVQSFPLAFVTEKLTGSGELMW 240
DB      181 TWTCTVLONQKVEFKIDIVLAFQKASSIVYKKEGEVQSFPLAFVTEKLTGSGELMW 240

QY      241 QABASSSKSWITFDLKNKEVSVKRVTDPKIQMGKULPLHLTLPOA 287
DB      241 QABASSSKSWITFDLKNKEVSVKRVISNPLRTTSSIIITTCIPSS 287

RESULT 52
5223394-9
; Patent No. 5223394
; APPLICANT: WALLNER, BARBARA
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
; LINKAGE SIGNAL SEQUENCE
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/335,688
; FILING DATE: 10-APR-1989
; SEQ ID NO:9:
; LENGTH: 295

Query Match      39.9%; Score 1363; DB 6; Length 295;
Best Local Similarity 99.3%; Pred. No. 2,2e-100;
Matches 266; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 NMRGVFRRLLLVLTQALLPAATQGNKVTLGKKGDTVELTCTASQKSIQFHWKNSNOIK 60

```


Db 1 MNRGPFPHLLLVQLALLPATQKQVVLGKGEVLELTCTASQKSIQFHMKNQIK 60
Qy 61 ILGNQSFPLTGPSTLNDNRADSRSLMNOGNFPLIKLKTEDDTYICEDQKEVQL 120
Db 61 ILGNQSFPLTGPSTLNDNRADSRSLMNOGNFPLIKLKTEDDTYICEDQKEVQL 120
Qy 121 LVPLGLTANSDPHLLQOGSLTLTLESPPGSSPVOCSPRGKNIOGKTLVSQLELODSG 180
Db 121 LVPLGLTANSDPHLLQOGSLTLTLESPPGSSPVOCSPRGKNIOGKTLVSQLELODSG 180
Qy 181 TWCTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGEVSEFPLAFTVEKLTJSGELMW 240
Db 181 TWCTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGEVSEFPLAFTVEKLTJSGELMW 240
Qy 241 QAERASSSKSWITFDLKNKEVSVKRVYQ 268
Db 241 QAERASSSKSWITFDLKNKEVSVKRVYR 268

RESULT 53

US-08-284-391B-33
Sequence 33, Application US/08284391B

Patent No. 5851828

GENERAL INFORMATION:

APPLICANT: Seed, Brian

APPLICANT: Banapour, Babak

APPLICANT: Romeo, Charles

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark & Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/284,391B

FILING DATE: 02-AUG-1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

TELEX:

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 254 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-284-391B-33

Query Match 39.2%; Score 1338.5; DB 2; Length 254;

Best Local Similarity 98.0%; Pred. No. 1.6e-98;
Matches 249; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

Qy 400 EPKSCDTHHC-----PELLGSPVFLFPPKPKDTLMISRTPEYTCVVVDVSHEDPEVKF 454
Db 1 EPKSCDTHHCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEYTCVVVDVSHEDPEVKF 60
Qy 455 NMVYDGVENNAKTPREEOYNSTYRVVSVLTVLHODMLNGKEYCKCKVSNKALPAPIEKT 514
Db 61 NMVYDGVENNAKTPREEOYNSTYRVVSVLTVLHODMLNGKEYCKCKVSNKALPAPIEKT 120
Qy 515 ISKAKGPREQVYTLPPSRDELTKQVSLTCLVKGFPSPDIAVEMESNGQPENNYKTP 574
Db 121 ISKAKGPREQVYTLPPSRDELTKQVSLTCLVKGFPSPDIAVEMESNGQPENNYKTP 180
Qy 575 PVLDSGSPFLYSKLTVDKSRWQGNVFGSGVMHEALHNYTQKSLSPQLDFTCAE 634
Db 181 PVLDSGSPFLYSKLTVDKSRWQGNVFGSGVMHEALHNYTQKSLSPQLDFTCAE 240
Qy 635 AQDGLDGLMTTDP 648
Db 241 AQDGLDGLMTTDP 254

RESULT 54

US-09-218-950-33
Sequence 33, Application US/09218950

Patent No. 6284240

GENERAL INFORMATION:

APPLICANT: Seed, Brian

APPLICANT: Banapour, Babak

APPLICANT: Romeo, Charles

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark & Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/218,950

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/284,391

FILING DATE: 02-AUG-1994

APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

TELEX:

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 254 amino acids

TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-33

Query Match 39.2%; Score 1338.5; DB 3; Length 254;
Best Local Similarity 98.0%; Pred. No. 1.6e-98;
Matches 249; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 400 EPKSCDKHTC-----PELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 454
DB 1 EPKSCDKHTCPCPCAPPELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
QY 455 NMVVDGVEVHNATKPREEQYNSTYRVSVLTVLHODMNGKRYCKVSNKALPAPIEKT 514
DB 61 NMVVDGVEVHNATKPREEQYNSTYRVSVLTVLHODMNGKRYCKVSNKALPAPIEKT 120
QY 515 ISKAKQPREPOVYTLPPSRDELITKNQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTP 574
DB 121 ISKAKQPREPOVYTLPPSRDELITKNQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTP 180
QY 575 PVLDSDGSPFLYSKLTVDKSRMOQGNVFCSCVMHEALHNHYTQKSLSLSPGQLDETCAE 634
DB 181 PVLDSDGSPFLYSKLTVDKSRMOQGNVFCSCVMHEALHNHYTQKSLSLSPGQLDETCAE 240
QY 635 AQDGLDGLMTTDP 648
DB 241 AQDGLDGLMTTDP 254

RESULT 55
US-09-313-942-8
Sequence 8, Application US/09313942

Patent No. 6472179
GENERAL INFORMATION:
APPLICANT: REGENERON PHARMACEUTICALS, INC.
TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING
TITLE OF INVENTION: AND USING
FILE REFERENCE: REG 203-A
CURRENT APPLICATION NUMBER: US/09/313,942
CURRENT FILING DATE: 1999-05-19
PRIOR APPLICATION NUMBER: 09/313,942
PRIOR FILING DATE: 1999-05-19
PRIOR APPLICATION NUMBER: 60/101,858
PRIOR FILING DATE: 1998-09-25
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 592
TYPE: PRT
ORGANISM: Homo sapiens
US-09-313-942-8

Query Match 37.3%; Score 1275; DB 4; Length 592;
Best Local Similarity 47.3%; Pred. No. 5.9e-93;
Matches 308; Conservative 53; Mismatches 160; Indels 130; Gaps 19;

QY 20 PAATQGNKVVILGKKGVLELTCTASQ-KKSIQPHMKNNOIKILGNQSFILTKGPSKIND 78
DB 26 PAQEVARGVLTSLPGDSVLTLCPEVEDENATVHM-----VLKKPA----- 66
QY 79 RADSRRLMDQGNFPLIILKULKIEDSTIYCE-----VEQKEVQLVYFGLT 126
DB 67 -AASHPRMAGMGRRLRLRSVOLHDSNGYSCYRAGRAGTVMHLLVDVPEEPDLSCFRKS 125
QY 127 ANSPDTHLLOQSULTLESPPGSPSVQCSPRGKNIQGGKTLVSQLELDGSGTWTCTV 186
DB 126 PLSN-----VVEWGRSTPSLTTKA-----VLLVRFQNSPADPQEP 165
QY 187 LQNKVVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVTEKLTGSGEL----- 238
DB 166 QYSGESKFFSCQLAVPEGDSFIVSMCVASSVSGSKSKTQTFQ---GCGILQDDPPANI 222

QY 229 -----W-----MQAERASSKSM--ITFDLKNKEVSKKATQDPKQLQMGKCLPLAHT 283
DB 223 TVTAAKRNPRMLSVTWQDPHNSWSSFYRLRFELKRAERSKTF-----TMMVKDLOHCV 278
QY 284 LPQALPOYAGSGLTLALAKTKLHQEVNLVVMRATOLQKULTCVWGTPSPKMLSLK 343
DB 279 IH-----DAMSLRH-----VYQLRA---QEEFGQGEWSEMSPEAMGTW 315
QY 344 LENK-----EAKYSKREKPPVNLNPEAGMQLCLSSGOVLTLESNIKVLPTWSTPVCAP 399
DB 316 TESRSPAPANEVS---TPQALITKNKDDNLLFRS-----ANATSLVQD-----AG 360
QY 400 EPKSCDKHTC-----PELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 454
DB 361 EPKSCDKHTCPCPCAPPELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 420
QY 455 NMVVDGVEVHNATKPREEQYNSTYRVSVLTVLHODMNGKRYCKVSNKALPAPIEKT 514
DB 421 NMVVDGVEVHNATKPREEQYNSTYRVSVLTVLHODMNGKRYCKVSNKALPAPIEKT 480
QY 515 ISKAKQPREPOVYTLPPSRDELITKNQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTP 574
DB 481 ISKAKQPREPOVYTLPPSRDELITKNQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTP 540
QY 575 PVLDSDGSPFLYSKLTVDKSRMOQGNVFCSCVMHEALHNHYTQKSLSLSPG 625
DB 541 PVLDSDGSPFLYSKLTVDKSRMOQGNVFCSCVMHEALHNHYTQKSLSLSPG 591

RESULT 56
US-09-499-846-2
Sequence 2, Application US/09499846

Patent No. 6656728
GENERAL INFORMATION:
APPLICANT: Kavanaugh et al.
TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
TITLE OF INVENTION: RECEPTOR-IMMUNOLOGICAL FUSION
FILE REFERENCE: 035784/195012 (5784-
CURRENT APPLICATION NUMBER: US/09/499,846
CURRENT FILING DATE: 2000-02-07
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FaastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 622
TYPE: PRT
ORGANISM: Homo sapiens
US-09-499-846-2

Query Match 36.4%; Score 1243; DB 4; Length 622;
Best Local Similarity 48.3%; Pred. No. 2.2e-90;
Matches 299; Conservative 47; Mismatches 119; Indels 154; Gaps 23;

QY 109 CEVEDQKEVOLLVFGEL-TANSDTHLLOQSULTLESPPGSPSVQCS--RSRGRK---- 161
DB 55 CRLRDVQGINWLRDQVLAESNRITGEVEVQ-DVPAADSGLYACVTSPPSGSDTTY 113
QY 162 ---NIOGKTLVSQLELDSTG-----WTCVNLQNKVVEFKIDIVL 202
DB 114 FSVNVSADALPSSDDDDDDSSSEKETDNTKRNPAVWT-----SPEMEKKHAV-- 166
QY 203 AFQKASSIYKKEGQVEFSPPLAFVE-KLTSQSGELMMQABRASSKSMITFDLKN-XE 260
DB 167 -----PAAKTIVKRCPSG-----TPNFTLKN-----LNGKE 194
QY 261 VSVKRVTDPKLQMG-----KULPLTLTQALPOYAGSGLTLALAKTKLHQEVNL-V 315
DB 195 FK-----PDHRIGYKRVATWSTIIMSVAP--SDKGVYTCIVENEGSINHVTQLDV 245
QY 316 VNRATQ--LQKUL-----TCEWGPSPKMLSLKLE----- 345
DB 246 VERSPHRPILOGLPANKTVALGSNVEFMCKVYSDQPHIOWLKHIEVNGSKIGPDNLPY 305
QY 346 ---NKEAKYSKREKPPVNLN-----PEAGMQLCLSDS----- 375

Db	306	VQLKTKAGVNTTQKXENVEVLHNRNVSPEDAGETVCLAGNSIGLSHHSANLTVLEALERPRA	365
Qy	376	---GQVLLBSNIKVLPFTWS--TPVPCPAPRPKSCDKTHTC-----PELLGGSYFLFPKP	426
Db	366	WMTSPVLYESRGELVPRGSGSP---GLOEPKSCDKTHTCPCPCPAPPELLGGSYFLFPKP	422
Qy	427	KQTLMTLSRPEVYCVVVDVSHDEPDKFMWYDGVENHNATTKRREQYNSTYRVSVLT	486
Db	423	KQTLMTLSRPEVYCVVVDVSHDEPDKFMWYDGVENHNATTKRREQYNSTYRVSVLT	482
Qy	487	VLHQDWLNGKEYYCKCKSNKALPAPLEIKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT	546
Db	483	VLHQDWLNGKEYYCKCKSNKALPAPLEIKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT	542
Qy	547	LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDYSRMQQGNVESCSV	606
Db	543	LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDYSRMQQGNVFCSV	602
Qy	607	MHEALHNHTQKSLSLSPG	625
Db	603	MHEALHNHTQKSLSLSPG	621

```

RESULT 57
US-09-590-656-2
; Sequence 2, Application US/09590656
; Patent No. 641932
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; APPLICANT: Borges, Luis G.
; APPLICANT: Fanslow, Iir, William C.
; TITLE OF INVENTION: TEK ANTACONISTS
; FILE REFERENCE: 2900-A
; CURRENT APPLICATION NUMBER: US/09/590,656
; CURRENT FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/137,889
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 704
; TYPE: PRF
; ORGANISM: Homo sapiens
; US-09-590-656-2

```

Query Match	36.2%	Score 1237	DB 4	Length 704
Best Local Similarity	81.4%	Pred. No. 7,9e-90		
Matches	240	Conservative	9	Mismatches 12; Indels 34; Gaps 5
QY	360	VLANBAGMOCLLSD-SGCVLLIESNI-KVLEPTMSTPVCPCAP-----	399	
Db	414	ILPDSGVMGVSVTYVAGMVEKPFNISVXKVLPK-----PLNAPNVIDGHNFAVINISSE	468	
QY	400	-----EPKSCDQTHHC-----PELLGGSPVFLFPKPKDQTLMISTRPEYTCVVDVSHSDP	450	
Db	469	PYFEPKPSKCDTHHCPCPCAPPELLGGSPVFLFPKPKDQTLMISTRPEYTCVVDVSHSDP	528	
QY	451	EVKKNMYYDGVGVHNNAKTKPREEOYNSTRYRVSVLTVLHODMLNGKEYCKKCVSNKALPAP	510	
Db	529	EVKKNMYYDGVGVHNNAKTKPREEOYNSTRYRVSVLTVLHODMLNGKEYCKKCVSNALPAP	588	
QY	511	IEKTSISKAGQPREQOYVTLPPSRDELTKNOVSLCLVKGFPSPDIAEWESNNGQPKNNY	570	
Db	589	IEKTSISKAGQPREQOYVTLPPSRDEMTKNOVSLCLVKGFPSPDIAEWESNNGQPKNNY	648	
QY	571	KTTTPPVLDSDGSFFLYSKLTYDKSRMOQGNVPSGSVMHEALHNHYTKSLSPG	625	
Db	649	KTTTPPVLDSDGSFFLYSKLTYDKSRMOQGNVPSGSVMHEALHNHYTKSLSPG	703	

RESULT 58
US-09-733-764-2

1 Patent No. 5231474
 2 Sentence 2, Application US/09733764
 3
 4 GENERAL INFORMATION:
 5
 6 APPLICANT: Cerretti, Douglas P.
 7 APPLICANT: Borges, Luis G.
 8 APPLICANT: Farnlow, Iii, William C.
 9 TITLE OF INVENTION: TEK ANTAGONISTS
 10 FILE REFERENCE: 2900-A
 11
 12 CURRENT APPLICATION NUMBER: US/09/733,764
 13 CURRENT FILING DATE: 2000-12-07
 14
 15 PRIOR APPLICATION NUMBER: 09/590,656
 16 PRIOR FILING DATE: 1999-06-07
 17
 18 NUMBER OF SEQ ID NOS: 2
 19
 20 SOFTWARE: PatentIn Ver. 2.0
 21
 22 SEQ ID NO 2
 23
 24 LENGTH: 704
 25
 26 TYPE: prt
 27
 28 ORGANISM: Homo sapiens
 29
 30 US-09-733-764-2

Query Match	36.2%	Score 1237;	DB 4	Length 704;
Best Local Similarity	81.4%	Pred. No. 7.9e-90;		
Matches 240;	Conservative 9;	Mismatches 12;	Indels 34;	Gaps 5.

Qy	360	VLANEAGMOOCLSD -SGOVLNESNI--KVLPTMTSTVPCAP-----	399
Db	414	ILPFDGVMVCSVNTVAGMEKPNINISKVLPK-----PLNAVINDGHNPAVINISSE	468
Qy	400	---EPKSCDKTHTC-----PELLGGSVFLPPEPKDXTLMSITPEVTCVVVDVSHEDP	450
Db	469	PYFGEPSKCDKTHTCPCPCPAPELLGGSVFLFPPEPKDXTLMSITPEVTCVVVDVSHEDP	528
Qy	451	EVKNNMVDGVEVNNATKPREEOYNSYRVSVLTJLHOMLNGKEYKCKVSKALPAP	510
Db	529	EVKNNMVDGVEVNNATKPREEOYNSYRVSVLTJLHOMLNGKEYKCKVSKALPAP	588
Qy	511	IEKTIISAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNQPENNY	570
Db	589	IEKTIISAKQPREPOVYTLPPSRDEMTKNOVSLTCLVKGFPSPDIAVEMESNQPENNY	648
Qy	571	KTPPVPVLDSDGSFLLYKSLTYDKSRMOOGNVFSCSVNHEALAHNYTQKSLSLSG	625
Db	649	KTPPVPVLDSDGSFLLYKSLTYDKSRMOOGNVFSCSVNHEALAHNYTQKSLSLSG	703

```

RESULT 59
US-09-499-846-6
: Sequence 6, Application US/09499846
: Patent No. 6656728
: GENERAL INFORMATION:
: APPLICANT: Kavanaugh et al.
: TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
: TITLE OF INVENTION: RECEPTOR-IMMUNOGLOBULIN FUSION
: FILE REFERENCE: 035784/195012 (5784-
: CURRENT APPLICATION NUMBER: US/09/499,846
: CURRENT FILING DATE: 2000-02-07
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PasteSeq for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 497
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-499-846-6

```

Qy	293	GSGLTTLTAATKATKLTG--EVLVLYMRAITOLQKL--TCEVWGSPISPKMLMSLKE--	345
	:	:	:
Db	118	GSINHTYQTDVVERSPHPILOAGLPANKTVALGSGNVEFMCKYVSDPPIHIOWLKHEVN	177
	:	:	:
Qy	346	-----NKEAKYSKREKPPWVLTN-----PEAGMMOCLLSDSGQVLLSNIK	385
	:	:	:

```

Db      178 GSKIGPDNLPIYVOLLKTAGVNTTDKEMEVHLRLNVSFEDAGEYTCCLAGNSIGLSHNS--- 234
      386 VLPTWST-----PVPCAP-----EPKSCDKTHTC-----PELLGGSVFL 420
Qy      235 ---AMLTVLEALBERPAVMTSPPLYLESGSGPGLQEPKSCDKTHTCPCAPPELLGGSVFL 231
Db      421 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKRREQDYNSTYR 480
Qy      292 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKRREQDYNSTYR 351
Db      481 VSVLTVLHODMNLNGEKYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELITKN 540
Qy      352 VSVLTVLHODMNLNGEKYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELITKN 411
Db      541 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSGSEFLLYSKLTVDKSRWQQGN 600
Qy      412 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSGSEFLLYSKLTVDKSRWQQGN 471
Db      601 VFSCSVHHEALHNHYTOKSLSLSPG 625
Qy      472 VFSCSVHHEALHNHYTOKSLSLSPG 496

RESULT 60
US-09-499-846-4
; Sequence 4, Application US/09499846
; Patent No. 6656728
; GENERAL INFORMATION:
; APPLICANT: Kavanagh et al.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
; FILE REFERENCE: 035784/195012 (5784-
; CURRENT APPLICATION NUMBER: US/09/499, 846
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-499-846-4

Query Match      36.2%; Score 1235.5; DB 4; Length 525;
Best Local Similarity 66.2%; Pred. No. 6.8e-90;
Matches 255; Conservative 18; Mismatches 54; Indels 58; Gaps 8;

```

```

Db      500 VFSCSVHHEALHNHYTOKSLSLSPG 524

RESULT 61
PCT-US95-03866-12
; Sequence 12, Application PC/TUS9503866
; GENERAL INFORMATION:
; APPLICANT: Cytomed, Inc. (all states except US)
; APPLICANT: Nocka, Karl (US only)
; APPLICANT: Lobell, Robert B (US only)
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,379
; FILING DATE: 28-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Cytomed/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-03866-12

Query Match      36.1%; Score 1232.5; DB 5; Length 424;
Best Local Similarity 88.9%; Pred. No. 8.6e-90;
Matches 232; Conservative 6; Mismatches 18; Indels 5; Gaps 1;

```

PCT-US95-03866-14
Sequence 14, Application PC/US9503866
GENERAL INFORMATION:
APPLICANT: Cytomed, Inc. (all states except US)
APPLICANT: Nocka, Karl (US only)
APPLICANT: Lobell, Robert B (US only)
TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/03866
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,379
FILING DATE: 28-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr, James F
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: Cytomed/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-03866-14
Query Match 36.1%; Score 1232.5; DB 5; Length 424;
Best Local Similarity 88.9%; Pred. No. 8.6e-90;
Matches 232; Conservative 6; Mismatches 18; Indels 5; Gaps 1;
QY CLTSDSGQVLLSNIKVLPTWSTVPVCPAPPEKSCDKHTC-----PELLGSPSVFLPP 424
DB 163 CVVSTSLSPKDSRVSYTKPFMLPVPVADPEPKSCDKHTCPCPAPPELLGSPSVFLPP 222
QY 425 KPQDTLMSRTPEVTCVVVDVSHEDPEVKFNWYDGVENNAKTRPREQYNSTYRVVSV 484
DB 223 KPQDTLMSRTPEVTCVVVDVSHEDPEVKFNWYDGVENNAKTRPREQYNSTYRVVSV 282
QY 485 LTVHOMLNKEKCKKYSNKAALPALEKTIISKAKGQPREQVYTLTPSRBELTKNQVSL 544
DB 283 LTVHOMLNKEKCKKYSNKAALPALEKTIISKAKGQPREQVYTLTPSRBELTKNQVSL 342
QY 545 TCLVKGFPYSDIAVEWESNGQPENNYKTTPEVLDSGFFLYSKLTVDYKSKWQQGNFSC 604
DB 343 TCLVKGFPYSDIAVEWESNGQPENNYKTTPEVLDSGFFLYSKLTVDYKSKWQQGNFSC 402
QY 605 SVMHEALHNYTKQSLSPG 625
DB 403 SVMHEALHNYTKQSLSPG 423
RESULT 63
US-08-157-101A-7
Sequence 7, Application US/08157101A
GENERAL INFORMATION:

APPLICANT: KURIHARA, TATSUYA
APPLICANT: MATSUKURA, SHIGEKAZU
APPLICANT: TSUBOKA, NOBUO
APPLICANT: ARIHA, KENJI
APPLICANT: NISHIHARA, TATSURO
TITLE OF INVENTION: ANTI-HBs ANTIBODY GENES AND EXPRESSION
TITLE OF INVENTION: PLASMIDS THEREFOR
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/157,101A
FILING DATE: 05-APR-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: TITUS, MARIANA K
REGISTRATION NUMBER: 35843
REFERENCE/DOCKET NUMBER: 9437/204199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3711
TELEFAX: 202-822-0944
TELEX: 6714627 CUCH
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-157-101A-7
Query Match 36.1%; Score 1232.5; DB 1; Length 459;
Best Local Similarity 46.7%; Pred. No. 9.7e-90;
Matches 290; Conservative 41; Mismatches 91; Indels 199; Gaps 21;
QY 25 GNRVILGKKGDTVELTCTAS--QKSIQFM-----KNSNQIKIL--GNQGSFL--TK 71
DB 17 GGGVV--QPRSLRLSCAAGFTFSSNSMHWVQAPCKGLEWAVILYDGNHKEFYADSVK 74
QY 72 GPKLNDRADRSRLMDQGNFPLIKLKLTEDSDTYICEVEDQKEVQLVFGLTANSDT 131
DB 75 GRFTIS--RDNSSKNTLY-----LEVKSLQEDTGVVYC--IRDQ-----TYGV----- 113
QY 132 HLQD--GQSLTLTLESPPGSSPSVQCSPPGKNIQGGKTLVSQLELDQSGTWCTVLQN 189
DB 114 HRDSDMGQGLLVYSSASTGSPVFLAPBSKSTSG--TALAGL----- 157
QY 190 QKVEFKIDIVLAFOKASIVYKKEGQVEFSFPPIAFVTEKLTGSGELMWQARASSK 249
DB 158 -----VKQYFPEPVTVS-----MNSGALASG-- 178
QY 250 SWITFLDKKEVSKVKTQDPKLOMGKGLPLHLTLPLALPOYVAGSGNLTALAEKTKLH 309
DB 179 -----VH--TFPAVL--QSSGLYLSLSVTVPPSSLG 206
QY 310 QEVNLVVMRATQLOKNLTCEVWGPTSPKMLSLKLEKKAIVSKREKPVVVLNLPDAGMWQ 369
DB 207 TQTYI-----CNV-----NHRKSNTRKVDKV----- 227
QY 370 CLTSDSGQVLLSNIKVLPTWSTVPVCPAPPEKSCDKHTC-----PELLGSPSVFLPP 424
DB 228 -----EPKSCDKHTCPCPAPPELLGSPSVFLPP 257

Qy	425	KPQDTLMI	STPEVTCV	VDVDSH	DEPEVFN	YVGV	VHNAKTRP	BEQVNS	YRVRVSV	484
Db	258	KPQDTLMI	STPEVTCV	VDVDSH	DEPEVFN	YVGV	VHNAKTRP	BEQVNS	YRVRVSV	317
Qy	485	LTVLHODM	LNGKEYK	CKVSNKAL	PAPIEKT	ISKAGOP	REPOVYTL	PPSRDEL	TKNOVSL	544
Db	318	LTVLHODM	LNGKEYK	CKVSNKAL	PAPIEKT	ISKAGOP	REPOVYTL	PPSRDEL	TKNOVSL	377
Qy	545	TCIVKGGY	PSPDIA	VEMESNGO	PENNYKTT	PVLDSGS	FPYLSK	TVDSK	MOQGNVFC	604
Db	378	TCIVKGGY	PSPDIA	VEMESNGO	PENNYKTT	PVLDSGS	FPYLSK	TVDSK	MOQGNVFC	437
Qy	605	SVNH	EALHNNHY	YTKSL	SLSPG	625				
Db	438	SVNH	EALHNNHY	YTKSL	SLSPG	458				

RESULT 64
US-08-397-411-7

```

1Sequence 7, Application US/08397411
2Patent No. 6129914
3
4GENERAL INFORMATION:
5APPLICANT: Weiner, George
6APPLICANT: Gingrich, Roger
7APPLICANT: Link, Brian
8APPLICANT: Teo, J. Yun
9TITLE OF INVENTION: Bispecific Antibody Effective to Treat
10TITLE OF INVENTION: B-Cell Lymphoma and Cell Line
11NUMBER OF SEQUENCES: 14
12CORRESPONDENCE ADDRESS:
13ADDRESS: Townsend and Townsend and Crew
14STREET: One Market Plaza, Steuart Tower, Suite 2000
15CITY: San Francisco
16STATE: California
17COUNTRY: USA
18ZIP: 94105
19
20COMPUTER READABLE FORM:
21MEDIUM TYPE: Floppy disk
22COMPUTER: IBM PC compatible
23OPERATING SYSTEM: PC-DOS/MS-DOS
24SOFTWARE: PatentIn Release #1.0, Version #1.25
25CURRENT APPLICATION DATA:
26APPLICATION NUMBER: US/08/397,411
27FILING DATE: 01-MAR-1995
28CLASSIFICATION: 424
29
30PRIOR APPLICATION DATA:
31APPLICATION NUMBER: US 07/859,583
32FILING DATE: 27-MAR-1992
33ATTORNEY/AGENT INFORMATION:
34NAME: Smith, William M.
35REGISTRATION NUMBER: 30,223
36REFERENCE/DOCKET NUMBER: 011823-004901
37TELECOMMUNICATION INFORMATION:
38TELEPHONE: 415-326-2400
39TELEFAX: 415-326-2422
40
41INFORMATION FOR SEQ ID NO: 7:
42SEQUENCE CHARACTERISTICS:
43LENGTH: 446 amino acids
44TYPE: amino acid
45STRANDEDNESS: single
46TOPOLOGY: linear
47
48MOLECULE TYPE: peptide
49
50US-08-397-411-7

```

Query Match	36.0%;	Score 1230.5;	DB 3;	Length 446;
-------------	--------	---------------	-------	-------------

Best Local Similarity 46.7%; Pred. No. 1.3e-89;
Matches 284; Conservative 36; Mismatches 103; Indels 185; Gaps 14.

QY 30 LGKGDVLTCTASQKSIQF--HWKNSNQIKILGNQGSFLTGP SKLNDRADSRSL- 86

Db 11 LVKPSETLSTCTVSGFSLTNYGVHWVRQSPGKGLEWIGVKWSGGSTEYNAAFISRLTIS 700

QY 87 - -WDQGNFPLIKNLKIEDSDTYICEVEDQKEVQLLVFGLTANS DTHLLQ--GQSLTLT 142

Db	71	KDTSKNQVSLKNTLSLTADTAIVYCC-----	ARRDYAMADWYGGGTLVT	113
OY	143	LESPGSSPSVQCHSPRGKNTGGGKTLSSVSGLELSDSGTWCTTVLQNGKKEFKIDIVVL		202
Db	114	VSSASTKQSPVFPPLAPSSKSTSGG-TAALGCL-		144
OY	203	AFQASSIVYKKEGEHOYFFSFPLAFTVEKLTGSGELMMQAEBAASSKSWITFDLKNKEVS		262
Db	145	-----VKDYFPPEPTVS-----	WNMGALTSQ-----	165
OY	263	VKRVYTDQPKLQMGKKLPLHLPLPOLPYQASGNTLALBACTGKLNHOEVLNVMRATOL		322
Db	166	-----VH-TFPAYL-QSSGILSSVTVVSSLSGTQYI-----		198
OY	323	QKNLTCEVWGPTSPKMLMLSLKLENKAYKSKREKELVWVLNPEAGMMQCLLSDSGVILSS		382
Db	199	-----CNV-----	NHKPSNTKKVDKY-----	214
OY	383	NIKVLPTWSTPEVCPAPKPSCKDKTHTCC-----	PELLGSPVFLPPPKKDTLMISRTPE	437
Db	215	-----EPKSCDKTHTCPAPKPELLGSPVFLPPPKKDTLMISRTPE		257
OY	438	VTCCVVVDVSHEDPEYKRWYVDGVEVNAKTKPRREOVNSTRVYVSLTVLHODWMLNGRE		497
Db	258	VTCCVVVDVSHEDPEYKRWYVDGVEVNAKTKPRREOVNSTRVYVSLTVLHODWMLNGRE		317
OY	498	YKCKVSNKALPAPIEKTISKAGQPREPOVYTLPPSRDELTAQNOVSLTCLVKGFPYSIDA		557
Db	318	YKCKVSNKALPAPIEKTISKAGQPREPOVYTLPPSRDELTAQNOVSLTCLVKGFPYSIDA		377
OY	558	VEMESNGQPENNYKTTTPVYLDSDGSFFLYSKLTVDKSKRQOGNVTSCSVMHEALHNHYQO		617
Db	378	VEMESNGQPENNYKTTTPVYLDSDGSFFLYSKLTVDKSKRQOGNVTSCSVMHEALHNHYQO		437
OY	618	KSLSLSPG 625		
Db	438	KSLSLSPG 445		

RESULT 65

US-08-776-511-2
Sequence 2, Application US/08776511
Patent No. 6153190
GENERAL INFORMATION:
APPLICANT: Young, Peter R.
APPLICANT: Erickson-Miller, Connie
TITLE OF INVENTION: Method for Obtaining Receptor Agonists
TITLE OF INVENTION: Antibodies
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation- Corporate
ADDRESSEE: Patents
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776,511
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jervie, Herbert H.
REGISTRATION NUMBER: 31,171
REFERENCE/DOCKET NUMBER: SBC P50349-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5015

```

; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 488 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-776-511-2

Query Match      35.9%; Score 1225.5; DB 3; Length 488;
Best Local Similarity 56.9%; Pred. No. 3.8e-89;
Matches 271; Conservative 36; Mismatches 92; Indels 77; Gaps 14;

QY 204 FOXASIIYKKEGGEVERSPFLAFVKEKLTGSGELIMQAEARASSKMTF--DKKNEV 261
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 35 FESKALILARGPRL-----LCFT-ERLEDLVCFMERASAGCPGNYSTYQLEDEPW 88

QY 262 SVKVTODPKLQMGK-----LPHLTLPLQALPOY--AGSNLTALAEAK 304
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 89 KLCLRHQAPTAGAVRFWCSLPTADTSGFVPLELRVTASGAPRHHVTHINEVLLDAP 148

QY 305 TG---KLHGEVNLVYMR-----ATOLQNLTCGEVWGPTSPKLMSLKENKAYSKREK 356
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 149 VGLVARLADSGHVLRLWLPPEETPMTHSHREYV-----DVSAGNGAGSVORVE- 197

QY 357 PWTVLNPEAGMOCCLSD--SGQVLESNIKV-----LPTMSTPVPCEAP----- 399
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 198 ---ILB---GRTECVSLMRGRTRYTFAVRAARMAEPGCGWASMSSEVSLTTSDDPI 251

QY 400 -----EPKSCDKTHTC-----PELIGSPVFLPPPKPDTLMISRTPEVTCVVVDVSHED 449
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 252 EGRGTEPESADKTHTCPCPCAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHED 311

QY 450 PEVKFNWYVDGEVYHNAKTKRREDOYNTYRVSVLYTLHODWLNKGEYKCKVSNKALPA 509
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 312 PEVKFNWYVDGEVYHNAKTKRREDOYNTYRVSVLYTLHODWLNKGEYKCKVSNKALPA 371

QY 510 PIETKISAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPRENN 569
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 372 PIETKISAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPRENN 431

QY 570 YKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHEALHNHYTQKSLSLSPG 625
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 432 YKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHEALHNHYTQKSLSLSPG 487

RESULT 66
US-09-499-846-12
; Sequence 12, Application US/09499846
; Patent No. 6656728
; GENERAL INFORMATION:
; APPLICANT: Kavanaugh et al.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
; FILE REFERENCE: 035784/195012 (5784-
; CURRENT APPLICATION NUMBER: US/09/499.846
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-499-846-12

Query Match      35.9%; Score 1225; DB 4; Length 488;
Best Local Similarity 67.3%; Pred. No. 4.2e-89;
Matches 253; Conservative 18; Mismatches 56; Indels 49; Gaps 8;

QY 293 GSGNLTALAEAKTGKLG--EVLNVRATOLQNL--TCEVWGPTSPKLMSLKLE-- 345
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 118 GSNHTYQLDVESPHRPIQAGLPANKTVALGSNVFEMCKVYSDPQHILMKHIEVN 177

```

```

QY 346 -----NKEAKYSKREKPVAVLN-----PEAGMOCCLSDSGQVLESNIK 385
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 GSKIGPNDLPVYOLTKTAGVNTTDKEMEVHLNRVSPEDAGEYTCLAGNSIGLSHHS--- 234

QY 386 VLPWTST-----PVPCEAP---EPKSCDKTHTC-----PELIGSPVFLPPPKPDT 429
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 235 ---AMLTVLEALERPAVMTSPFLYEPKSCDKTHTCPCPCAPLEGGSPSVFLFPKPKDT 291

QY 430 LMTSRTPEVTCVVVDVSHEDPEVKFNWYVDGEVYHNAKTKRREDOYNTYRVSVLYTLH 489
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 292 LMTSRTPEVTCVVVDVSHEDPEVKFNWYVDGEVYHNAKTKRREDOYNTYRVSVLYTLH 351

QY 490 QDMLNGEYKCKVSNKALPAPIETKISAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 549
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 352 QDMLNGEYKCKVSNKALPAPIETKISAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 411

QY 550 GFYPSDIAVEMESNGQPRENNYKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHE 609
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 412 GFYPSDIAVEMESNGQPRENNYKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHE 471

QY 610 ALHNHYTQKSLSLSPG 625
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 472 ALHNHYTQKSLSLSPG 487

RESULT 67
US-08-227-496C-15
; Sequence 15, Application US/08227496C
; Patent No. 6130202
; GENERAL INFORMATION:
; APPLICANT: Greve, Jeffrey M.
; APPLICANT: McClelland, Alan
; TITLE OF INVENTION: Multimeric Forms of Human
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bayer Corporation
; STREET: 400 Morgan Lane
; CITY: West Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06516
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette, 1.44 Mb storage
; COMPUTER: Dell Optiplex GX1
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect 8.0 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,496C
; FILING DATE: 04/14/94
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/903,069
; FILING DATE: 06/22/92
; APPLICATION NUMBER: 07/704,984
; FILING DATE: 05/24/91
; APPLICATION NUMBER: 07/556,238
; FILING DATE: 07/20/90
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara A. Shime1
; REGISTRATION NUMBER: 29,862
; REFERENCE/DOCKET NUMBER: MTI 214.2C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 812-5492
; TELEFAX: (203) 812-5492
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 680 amino acid residues
;   TYPE: amino acids
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: no
; HYPOTHETICAL: no

```

```

; FRAGMENT TYPE: complete sequence
;
; NAME/KEY: t1CAM(185)/19G fusion protein
; OTHER INFORMATION: amino acid residues 1-453 =
; OTHER INFORMATION: t1CAM(453); amino acid residues 454-680 = amino
; OTHER INFORMATION: acid residues 216-442 of human Ig1 heavy chain
US-08-227-496C-15

```

```

Query Match      35.8%; Score 1223.5; DB 3; Length 680;
Best Local Similarity 50.5%; Pred. No. 8.9e-89;
Matches 285; Conservative 44; Mismatches 114; Indels 121; Gaps 18;

```

```

QY 99 LKIEDSTYICEVEDQKEEVLVFGI--TANSDTHLQGSGLTLIESPPSGSSPQCK 156
DB 200 LEVDTCGTVCSDID-----GLPVSSEAQVHLALSDQRL-----NPTV--- 236
QY 157 SPFGKNIQCGKTLVSQLELDSDGT--WTCTVLQNKQKVEFKIDIVLAFQKASIVYKK 214
DB 237 -TGNDSFSAKA-SVS-VTAEDEGTQRLTCVILGNOSQTLQTVITISFPANVILTKP 293
QY 215 EGEQVEFSPFLAFTVEKLTGSGELMWQAEKASSKSWITDLYNKEYSVK-----RVT 267
DB 294 EVSE-----GTEVTYVCEAHPPAKYT 314
QY 268 QD--PKLQMGKKLPLHLTLPOALPOYAG--SGNLTLEAKTGKLEHVNVLVMMRATOL 332
DB 315 LNVGPAOPLGPRKOL--LKATPEDNGRSFSCATLEVAQOLHKQOTRELVLVIGPRL 371
QY 323 -----QKLTCEVWGPTSPKLMLSLKLENKAKVSKREKPVWVLNPEAGM 367
DB 372 DERDCEGNWTWPNESQOTPCQAMGNPLPELK-CLNKGTFPLPIG--ESETVTRDLEGT 427
QY 368 WQC-LISDSQVLESNIKLPFWSTPVPCEPAPBPSCKDTHTC-----PELLGGPSVFL 421
DB 428 YLCRASTQGEVTRKVTNVNL-----SPRYEDKTHTCPPCAPBELLGGPSVFL 475
QY 422 FPKPKQDTLMISRTPEVTCVVVDVSHEDPEVKFMWYDGVENNAKTKPREEOYNSTRY 481
DB 476 FPKPKQDTLMISRTPEVTCVVVDVSHEDPEVKFMWYDGVENNAKTKPREEOYNSTRY 535
QY 482 VSVLTVLHQMMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQ 541
DB 536 VSVLTVLHQMMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQ 595
QY 542 VSVLTVLHQMMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQ 601
DB 596 VSVLTVLHQMMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQ 655
QY 602 FSCSVMEALHNHYTOKSLSLSPG 625
DB 656 FSCSVMEALHNHYTOKSLSLSPG 679

```

RESULT 68

```

US-07-934-373C-22
; Sequence 22: Application US/07934373C
; Patent No. 5821337

```

```

; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA

```

```

; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,373C
; FILING DATE: 21-Aug-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-07-934-373C-22

```

```

Query Match      35.8%; Score 1222.5; DB 2; Length 454;
Best Local Similarity 46.8%; Pred. No. 5.9e-89;
Matches 285; Conservative 35; Mismatches 110; Indels 179; Gaps 15;

```

```

QY 30 LGKKGDTVELTCTASQKKSIOF--HWKNSQIKILGNQCSFLTK-SPSKLNDPDRSRSL 86
DB 11 LKPGASVAKISCKTSQTYFTETMHMMKSHGKSLWEMIGFNPKNQSGSHNQRFMDKATL 70
QY 87 ---WDQGNFPLIKIKIEDSTYICEVEDQKEEVLVFGITANSPTHLQ--GQSLTL 141
DB 71 AVDKSTSTAYMELRSLTSEDGIIYC-----ARWRGLNYGFDVRYFDVWGAGTIV 120
QY 142 TLESPPGSSPVQOCSPRKQNIQCGKTLVSQLELDSDGTWCTVLQNKQKVEFKIDIV 201
DB 121 TVSSASTKGPVPLPSSKSTSG-TAALGCL----- 152
QY 202 LAFQKASIVYKKEGQVEFPLAFTVEKLTGSGELMWQAEKASSKSWITFDLYNKEY 261
DB 153 -----VDYFPEPVTVS-----WNSGLTSG----- 173
QY 262 SVKRVTDQPKLQMGKKLPLHLTLPOALPOYAGSGNLTLEAKTGKLEHVNVLVMMRATO 321
DB 174 -----VH-TFPAYL--QSSGLYSLSVTVVPSSSLGTQTYI----- 206
QY 322 LQKNLTCGEVWGPTSPKLMLSLKLENKAKVSKREKPVWVLNPEAGMQLSDSGVLL 381
DB 207 -----CNV-----NHKPSNTKYDKKV----- 222
QY 382 SNIKVLPTWSTVPCPAPBPSCKDTHTC-----PELLGSPSVFLPPPKQDTLMISRT 436
DB 223 -----EPKSDCKTHTCPPCAPBELLGGPSVFLPPPKQDTLMISRT 264
QY 437 EYTCVVVDVSHEDPEVKFMWYDGVENNAKTKPREEOYNSTRYVSVLTVLHQMMLNGK 496
DB 265 EYTCVVVDVSHEDPEVKFMWYDGVENNAKTKPREEOYNSTRYVSVLTVLHQMMLNGK 324
QY 497 EYCKCKSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVGYFPSDI 556
DB 325 EYCKCKSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVGYFPSDI 384
QY 557 AVEMESNGQPENNYKTTTPVLDSGDFLYSKLTVDKSRWQGNVSCSVMEALHNHYT 616
DB 385 AVEMESNGQPENNYKTTTPVLDSGDFLYSKLTVDKSRWQGNVSCSVMEALHNHYT 444
QY 617 QKSLSLSPG 625
DB 445 QKSLSLSPG 453

```



```

1      RESULT 69
2      US-08-437-642B--22
3      / Sequence 22, Application US/08437642B
4      / Patent No. 6054297
5      / GENERAL INFORMATION:
6      / APPLICANT: Paul J. Carter
7      / APPLICANT: Leonard G. Presta
8      / TITLE OF INVENTION: Immunoglobulin Variants
9      / NUMBER OF SEQUENCES: 47
10     / CORRESPONDENCE ADDRESS:
11     / ADDRESSEE: Genentech, Inc.
12     / STREET: 1 DNA Way
13     / CITY: South San Francisco
14     / STATE: California
15     / COUNTRY: USA
16     / ZIP: 94080
17
18     / COMPUTER READABLE FORM:
19     / MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
20     / COMPUTER: IBM PC compatible
21     / OPERATING SYSTEM: PC-DOS/MS-DOS
22     / SOFTWARE: Winpatln (Genentech)
23
24     / CURRENT APPLICATION DATA:
25     / APPLICATION NUMBER: US/08/437, 642B
26     / FILING DATE: 09-May-1995
27     / CLASSIFICATION: 530
28     / PRIOR APPLICATION DATA:
29     / APPLICATION NUMBER: 07/934373
30     / FILING DATE: 21-AUG-1992
31     / PRIOR APPLICATION DATA:
32     / APPLICATION NUMBER: 08/146206
33     / FILING DATE: 17-NOV-1993
34     / PRIOR APPLICATION DATA:
35     / APPLICATION NUMBER: PCT/US92/05126
36     / FILING DATE: 15-JUN-1992
37     / PRIOR APPLICATION DATA:
38     / APPLICATION NUMBER: 07/715272
39     / FILING DATE: 14-JUN-1991
40     / ATTORNEY/AGENT INFORMATION:
41     / NAME: Lee, Wendy M.
42     / REGISTRATION NUMBER: 40,378
43     / REFERENCE/DOCKET NUMBER: P0709P2C1
44     / TELECOMMUNICATION INFORMATION:
45     / TELEPHONE: 650/225-1994
46     / TELEFAX: 650/952-9881
47     / INFORMATION FOR SEQ ID NO: 22:
48     / SEQUENCE CHARACTERISTICS:
49     / LENGTH: 454 amino acids
50     / TYPE: Amino Acid
51     / TOPOLOGY: Linear
52
53     / US-08-437-642B--22
54
55     / Query Match      35.8%; Score 1222.5; DB 3; Length 454;
56     / Beec Local Similarity 46.8%; Pred. 5.9e-89;
57     / Matches 285; Conservative 35; Mismatches 110; Indels 179; Gaps 15.
58
59     /
60     /
61     /
62     /
63     /
64     /
65     /
66     /
67     /
68     /
69     /
70     /
71     /
72     /
73     /
74     /
75     /
76     /
77     /
78     /
79     /
80     /
81     /
82     /
83     /
84     /
85     /
86     /
87     /
88     /
89     /
90     /
91     /
92     /
93     /
94     /
95     /
96     /
97     /
98     /
99     /
100    /
101    /
102    /
103    /
104    /
105    /
106    /
107    /
108    /
109    /
110    /
111    /
112    /
113    /
114    /
115    /
116    /
117    /
118    /
119    /
120    /
121    /
122    /
123    /
124    /
125    /
126    /
127    /
128    /
129    /
130    /
131    /
132    /
133    /
134    /
135    /
136    /
137    /
138    /
139    /
140    /
141    /
142    /
143    /
144    /
145    /
146    /
147    /
148    /
149    /
150    /
151    /
152    /
153    /
154    /
155    /
156    /
157    /
158    /
159    /
160    /
161    /
162    /
163    /
164    /
165    /
166    /
167    /
168    /
169    /
170    /
171    /
172    /
173    /
174    /
175    /
176    /
177    /
178    /
179    /
180    /
181    /
182    /
183    /
184    /
185    /
186    /
187    /
188    /
189    /
190    /
191    /
192    /
193    /
194    /
195    /
196    /
197    /
198    /
199    /
200    /
201    /
202    /
203    /
204    /
205    /
206    /
207    /
208    /
209    /
210    /
211    /
212    /
213    /
214    /
215    /
216    /
217    /
218    /
219    /
220    /
221    /
222    /
223    /
224    /
225    /
226    /
227    /
228    /
229    /
230    /
231    /
232    /
233    /
234    /
235    /
236    /
237    /
238    /
239    /
240    /
241    /
242    /
243    /
244    /
245    /
246    /
247    /
248    /
249    /
250    /
251    /
252    /
253    /
254    /
255    /
256    /
257    /
258    /
259    /
260    /
261    /
262    /
263    /
264    /
265    /
266    /
267    /
268    /
269    /
270    /
271    /
272    /
273    /
274    /
275    /
276    /
277    /
278    /
279    /
280    /
281    /
282    /
283    /
284    /
285    /
286    /
287    /
288    /
289    /
290    /
291    /
292    /
293    /
294    /
295    /
296    /
297    /
298    /
299    /
300    /
301    /
302    /
303    /
304    /
305    /
306    /
307    /
308    /
309    /
310    /
311    /
312    /
313    /
314    /
315    /
316    /
317    /
318    /
319    /
320    /
321    /
322    /
323    /
324    /
325    /
326    /
327    /
328    /
329    /
330    /
331    /
332    /
333    /
334    /
335    /
336    /
337    /
338    /
339    /
340    /
341    /
342    /
343    /
344    /
345    /
346    /
347    /
348    /
349    /
350    /
351    /
352    /
353    /
354    /
355    /
356    /
357    /
358    /
359    /
360    /
361    /
362    /
363    /
364    /
365    /
366    /
367    /
368    /
369    /
370    /
371    /
372    /
373    /
374    /
375    /
376    /
377    /
378    /
379    /
380    /
381    /
382    /
383    /
384    /
385    /
386    /
387    /
388    /
389    /
390    /
391    /
392    /
393    /
394    /
395    /
396    /
397    /
398    /
399    /
400    /
401    /
402    /
403    /
404    /
405    /
406    /
407    /
408    /
409    /
410    /
411    /
412    /
413    /
414    /
415    /
416    /
417    /
418    /
419    /
420    /
421    /
422    /
423    /
424    /
425    /
426    /
427    /
428    /
429    /
430    /
431    /
432    /
433    /
434    /
435    /
436    /
437    /
438    /
439    /
440    /
441    /
442    /
443    /
444    /
445    /
446    /
447    /
448    /
449    /
450    /
451    /
452    /
453    /
454    /
455    /
456    /
457    /
458    /
459    /
460    /
461    /
462    /
463    /
464    /
465    /
466    /
467    /
468    /
469    /
470    /
471    /
472    /
473    /
474    /
475    /
476    /
477    /
478    /
479    /
480    /
481    /
482    /
483    /
484    /
485    /
486    /
487    /
488    /
489    /
490    /
491    /
492    /
493    /
494    /
495    /
496    /
497    /
498    /
499    /
500    /
501    /
502    /
503    /
504    /
505    /
506    /
507    /
508    /
509    /
510    /
511    /
512    /
513    /
514    /
515    /
516    /
517    /
518    /
519    /
520    /
521    /
522    /
523    /
524    /
525    /
526    /
527    /
528    /
529    /
530    /
531    /
532    /
533    /
534    /
535    /
536    /
537    /
538    /
539    /
540    /
541    /
542    /
543    /
544    /
545    /
546    /
547    /
548    /
549    /
550    /
551    /
552    /
553    /
554    /
555    /
556    /
557    /
558    /
559    /
560    /
561    /
562    /
563    /
564    /
565    /
566    /
567    /
568    /
569    /
570    /
571    /
572    /
573    /
574    /
575    /
576    /
577    /
578    /
579    /
580    /
581    /
582    /
583    /
584    /
585    /
586    /
587    /
588    /
589    /
590    /
591    /
592    /
593    /
594    /
595    /
596    /
597    /
598    /
599    /
600    /
601    /
602    /
603    /
604    /
605    /
606    /
607    /
608    /
609   
```

[illegible]

```

QY 30 LGKKGDTVELTCTAASQKSIQF--HWKNSNQIKILGNQGSFLTK-GPSKLNDRADSRRL 86
DB 11 LVKPGASVKISCKTSGYTFEYTMHMKQSHGKSLMIWGFNPKNGSSSHNQRFMDKATL 70
QY 87 ---WDQGNFPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFHLLQ--GQSLTL 141
DB 71 AVDKSTSTAYMELRSLTSEDSGIYYC-----ARMGGLNYGDFVRFDWGAGTIV 120
QY 142 TLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWCTVLQNKQVEFKIDIYV 201
DB 121 TVSSASTKPSVFPPLAPSSKSTSGG-TAALGCL----- 152
QY 202 LAFQKASSIVYKKEGEVEFSFPLAFVEKLTGSGELMWQABRASSSKMTFDLKNKEY 261
DB 153 -----VKDYFPEPVTVS-----NMSGALITSG----- 173
QY 262 SVKRYTQDPKLGKMLPLHLTPQALPOYAGSGNTLALAEKTKLHDEVNILVMWRATQ 321
DB 174 -----VH-TFPAVL-QSSGLYSLSVTVTPSSSLGTQTYI----- 206
QY 322 LQKNLTCEVWGPTSPKMLMSLKLENKAKVSKREKPVWVLPBAGMWQCILSDSGVLLH 381
DB 207 -----CNV-----NHKPSNTKYDKKV----- 222
QY 382 SNIKVLPTWSTPVPCEAPPEKSCDKTHTC-----PELLGSPSVFLFPPKXDTLMTSRP 436
DB 223 -----EPKSCDKTHTCPPCAPPELLGSPSVFLFPPKXDTLMTSRP 264
QY 437 EYTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQVNSTYRVVSVLTVLHQDMLNGK 496
DB 265 EYTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQVNSTYRVVSVLTVLHQDMLNGK 324
QY 497 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 556
DB 325 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 384
QY 557 AVEWESNGQPENNYKTTTPVLDSGSEFLYSKLTVDKSRWQGNVSCSMHEALHNHYT 616
DB 385 AVEWESNGQPENNYKTTTPVLDSGSEFLYSKLTVDKSRWQGNVSCSMHEALHNHYT 444
QY 617 QKSLSLSPG 625
DB 445 QKSLSLSPG 453

```

RESULT 71
 US-09-705-686-22
 ; Sequence 22, Application US/09705686
 ; Patent No. 6639055
 ;
 ; GENERAL INFORMATION:
 ; APPLICANT: Carter, Paul J.
 ; Presta, Leonard G.
 ; TITLE OF INVENTION: Method for Making Humanized Antibodies
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 1 DNA Way
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Minipain (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/705,686
 ; FILING DATE: 02-No. 6639055-2000
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/146206

```

; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P1D3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9681
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-705-686-22

Query Match      35.8%; Score 1222.5; DB 4; Length 454;
Best Local Similarity 46.8%; Pred. No. 5,9e-89;
Matches 285; Conservative 35; Mismatches 110; Indels 179; Gaps 15;

QY 30 LGKKGDTVELTCTAASQKSIQF--HWKNSNQIKILGNQGSFLTK-GPSKLNDRADSRRL 86
DB 11 LVKPGASVKISCKTSGYTFEYTMHMKQSHGKSLMIWGFNPKNGSSSHNQRFMDKATL 70
QY 87 ---WDQGNFPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFHLLQ--GQSLTL 141
DB 71 AVDKSTSTAYMELRSLTSEDSGIYYC-----ARMGGLNYGDFVRFDWGAGTIV 120
QY 142 TLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWCTVLQNKQVEFKIDIYV 201
DB 121 TVSSASTKPSVFPPLAPSSKSTSGG-TAALGCL----- 152
QY 202 LAFQKASSIVYKKEGEVEFSFPLAFVEKLTGSGELMWQABRASSSKMTFDLKNKEY 261
DB 153 -----VKDYFPEPVTVS-----NMSGALITSG----- 173
QY 262 SVKRYTQDPKLGKMLPLHLTPQALPOYAGSGNTLALAEKTKLHDEVNILVMWRATQ 321
DB 174 -----VH-TFPAVL-QSSGLYSLSVTVTPSSSLGTQTYI----- 206
QY 322 LQKNLTCEVWGPTSPKMLMSLKLENKAKVSKREKPVWVLPBAGMWQCILSDSGVLLH 381
DB 207 -----CNV-----NHKPSNTKYDKKV----- 222
QY 382 SNIKVLPTWSTPVPCEAPPEKSCDKTHTC-----PELLGSPSVFLFPPKXDTLMTSRP 436
DB 223 -----EPKSCDKTHTCPPCAPPELLGSPSVFLFPPKXDTLMTSRP 264
QY 437 EYTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQVNSTYRVVSVLTVLHQDMLNGK 496
DB 265 EYTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQVNSTYRVVSVLTVLHQDMLNGK 324
QY 497 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 556
DB 325 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 384
QY 557 AVEWESNGQPENNYKTTTPVLDSGSEFLYSKLTVDKSRWQGNVSCSMHEALHNHYT 616
DB 385 AVEWESNGQPENNYKTTTPVLDSGSEFLYSKLTVDKSRWQGNVSCSMHEALHNHYT 444
QY 617 QKSLSLSPG 625
DB 445 QKSLSLSPG 453

```

RESULT 72
 PCT-US93-07832-22
 ; Sequence 22, Application PC/TUS9307832
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; TITLE OF INVENTION: Immunoglobulin Variants

```

: NUMBER OF SEQUENCES: 40
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/07832
: FILING DATE: 19930820
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/715272
: FILING DATE: 14-JUN-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/05126
: FILING DATE: 15-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/934373
: FILING DATE: 21-AUG-1992
: ATTORNEY/AGENT INFORMATION:
: NAME:
: REGISTRATION NUMBER:
: REFERENCE/DOCKET NUMBER: 709P2PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE:
: TELEFAX: 415/952-9881
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 454 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: PCT-US93-07832-22

Query Match      35.8%; Score 1222.5; DB 5; Length 454;
Best Local Similarity 46.8%; Pred. No. 5.9e-89;
Matches 285; Conservative 35; Mismatches 110; Indels 179; Gaps 15;

QY 30 LGKGGTVLCTASQKSIQF--HWKNSNOIKILGNGSFLTK-GPSKINDRADRSRL 86
DB 11 LVKGGASVYKISCKTSGYFTFTYTHMMKQSHGKSLKEMIGFNPKNKGSSHQRFMDKATL 70
QY 87 ---WDQGNFPLIKLKIEDSDTYICEVEDQKEVQLLVPLGLTANSPTHLQ--GQSILTL 141
DB 71 AVDKSTSTAVYMLNSLTSEDSGIYYC-----ARWRGLNYGFDRYFDVWGAGTTV 120
QY 142 TLSEPPGSSPVQCCSPRGNIOGGKTLVSQLELDGSGTWTCTVLQNKVFEKIDIV 201
DB 121 TVSSASTGSPVFLAPSSKSTSG-TALAGL----- 152
QY 202 LAFQASSIVYKKEGQVEFSFPLAFVTEKLTSGSELMWAERASSKSKWTFDLKKEV 261
DB 153 ---VKDYFPEPVTVS-----WNSGALTSG----- 173
QY 262 SVKAVTQDPKIQMGKKLPLHLTLTQALPQYAGSGNLTALEAKGKLGHOENVLVMRATQ 321
DB 174 ---VH--TFPAVL--QSSGLYSLSSVVTPSSSLGTQTYI----- 206
QY 322 LQKNLTCEVMGPTSPKMLSLKLENKEAKVSKREKPVVNLNPEAGMOCLLSDSGQVLE 381
DB 207 ---CNV-----NHKSNTKVDKKV----- 222
QY 382 SNIKVLTWSTPVPCPAPEPKSCDKTHTC-----PELLGSPVFLFPKPKDTIMISTP 436
DB 223 ---EPKSCDKTHTCPCPAPELLGSPVFLFPKPKDTIMISTP 264

```

```

QY 437 EVTCVVVDVSHEDPEVKFNKYVDGVEVHNAKTKRPREQYNSTYRVSVLTITLVQDWLNGK 496
DB 265 EVTCVVVDVSHEDPEVKFNKYVDGVEVHNAKTKRPREQYNSTYRVSVLTITLVQDWLNGK 324
QY 497 EYKCKSNKKLPAPIETKISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 556
DB 325 EYKCKSNKKLPAPIETKISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 384
QY 557 AVESWESNGPENNYKTTPPLDSDGSPFLYSKLTVDKSRMQQCVFSCSVWHEALHNHYT 616
DB 385 AVESWESNGPENNYKTTPPLDSDGSPFLYSKLTVDKSRMQQCVFSCSVWHEALHNHYT 444
QY 617 QKSLSLSPG 625
DB 445 QKSLSLSPG 453

RESULT 73
US-09-740-002-25
: Sequence 25, Application US/09740002
: Patent No. 6537809
: GENERAL INFORMATION:
: APPLICANT: BRAMS, PETER
: TITLE OF INVENTION: NEUTRALIZING HIGH AFFINITY HUMAN MONOCLONAL ANTIBODIES
: TITLE OF INVENTION: SPECIFIC TO RSV F-PROTEIN AND METHODS FOR THEIR
: TITLE OF INVENTION: MANUFACTURE AND THERAPEUTIC USE THEREOF
: FILE REFERENCE: 037003-027579
: CURRENT APPLICATION NUMBER: US/09/740,002
: PRIOR FILING DATE: 2000-12-20
: PRIOR APPLICATION NUMBER: 09/335,697
: PRIOR FILING DATE: 1999-06-18
: PRIOR APPLICATION NUMBER: 08/488,376
: PRIOR FILING DATE: 1995-06-07
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 25
: LENGTH: 475
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-740-002-25

Query Match      35.8%; Score 1222.5; DB 4; Length 475;
Best Local Similarity 45.6%; Pred. No. 6.4e-89;
Matches 293; Conservative 35; Mismatches 111; Indels 203; Gaps 17;

QY 10 LLLVQLALPLPATQGNKVVLGKGGTVLCTAS-----QKSIQPHWK 54
DB 10 LVAVATRVLSQVQLQESGPVYVVKPTETLTCTVSGSLSPNRMGVWIRQPPGKALEW- 68
QY 55 NSNOIKILGN---QGSFLTKGPSKINDRADRSRLMDQGNFPLIKLKIEDSDTYIC 109
DB 69 ---LGNIFSSDEKFSFSLKSLRLTTSQDTSRS-----QVLSLTNVDPVDTATYYC 116
QY 110 EVEDQKEVQLLVGLTANSDTHL--LQGSILTLTLESPPSSPVQCCSPRGNIOGGK 168
DB 117 ---ARVGLDINAYLYLDYDWGQGLTVVSSASTKGSVPFLAPSSKSTSG-T 167
QY 169 LSVSQLELDGSGTWTCTVLQNKVFEKIDIVLAFQASSIVYKKEGQVEFSFPLAF 228
DB 168 LALGCL-----VKDYFPEPVTVS----- 183
QY 229 VEKLTSGSELMWAERASSKSKWTFDLKKEVSVKAVTQDPKIQMGKKLPLHLTLTQAL 288
DB 184 VS-----WNSGALTSG-----VH--TFPAVL 202
QY 289 PQYAGSGNLTALEAKGKLGHOENVLVMRATQIQKNLTCEVMGPTSPKMLSLKLENKE 348
DB 203 ---QSSGLYSLSSVVTPSSSLGTQTYI-----CNV----- 230
QY 349 AKVSKREKPVVNLNPEAGMOCLLSDSGQVLESNIKVLTWSTPVPCPAPEPKSCDKTH 408
DB 231 ---NHKP-----SNTKV-----DKKAEPSKCDKTH 252

```

QY 409 TC-----PELLGGPSVFLPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNRYVDGVEV 463
DB 253 TCPCCAPPELLGGPSVFLPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNRYVDGVEV 312
QY 464 HNAKTKREDOYNSYVAVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPR 523
DB 313 HNAKTKREDOYNSYVAVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPR 372
QY 524 EPQYVITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGOPENNYKTTPVLDSDGSF 583
DB 373 EPQYVITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGOPENNYKTTPVLDSDGSF 432
QY 584 FLYSKLTVDKSRMGOQGVFSCSVMHAEALHNHYTQKSLSLSPG 625
DB 433 FLYSKLTVDKSRMGOQGVFSCSVMHAEALHNHYTQKSLSLSPG 474

RESULT 74
US-09-049-672A-4
Sequence 4, Application US/09049672A
Patent No. 6135941
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Tang, Y. Tom
APPLICANT: Yue, Henry
APPLICANT: Au-Young, Janice
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/049,672A
FILING DATE: HEREMITH
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0497 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 473 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PANTUT01
CLONE: 1513264
US-09-049-672A-4

Query March 35.8%; Score 1221.5; DB 3; Length 473;
Best Local Similarity 45.3%; Pred. No. 7.6e-89;

Matches 288; Conservative 40; Mismatches 125; Indels 183; Gaps 15;
QY 8 RHLLVLQALLP-----AATQGNKVVLGKGDYBELTCTAS--QKSIQIFWKRNSQI 59
DB 2 KHLMFELLVAAPRWYLSQVLOESGPGLVKSEETLSLTCVAVSGSITSGGYWMSWIRP 61
QY 60 KILGNO--GSPFLTKGPSKINDRASRRSL---WDQGNFLIITKNLKIETSDTYICVEQD 114
DB 62 PGKGLIEWITGIYISGSLTYNPSLSKRVITISVTSKQFSLKLSVTAADTAIYYCARD- 120
QY 115 KEEOVLVEGLTANSDTHLLQOSLTLLLESPGSSPVQCRSPRKNIIQGGKTLVSQI 174
DB 121 -----VGLGANGYGMVWGQGLTVTSASTKGPSPFLAPSSKSTSG--TALGCL 171
QY 175 ELQDSGTWTCYVLQNKQKIEFKIDIVLAFOKASSIVYKKEGQVEFSFPLATVEKLG 234
DB 172 -----VQDFPEEPTVS----- 183
QY 235 SGEIWMQAEARASSSKSWITFDLKNKEVSRYTQDPKLQMGKLPALITLPQALPYAGS 294
DB 184 -----INSGALISG-----VH-TFPVAVL--QSSGL 205
QY 295 GNITLALFAKTKGLHQEVNLVVMRATQLOKNLTCEYWGPTSPKMLSLKLNKAARVSKR 354
DB 206 YSLSSVTVTPSSSLGTYI-----CNY-----NKKPSTKV 237
QY 355 EKPVWVLANPEAGMOCCLSDSGOVLIESNIKVLPTWSTVPCPAPRPSGCDKHTHC---- 410
DB 238 DKRV-----EPKSCDKHTHTPPCP 256
QY 411 -PELLGGPSVFLPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNRYVDGVEVHNAKTK 469
DB 257 APPELLGGPSVFLPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNRYVDGVEVHNAKTK 316
QY 470 PREQYNSTYRVAVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREQVYT 529
DB 317 PREQYNSTYRVAVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREQVYT 376
QY 530 LPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGOPENNYKTTPVLDSDGSFLYSKL 589
DB 377 LPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGOPENNYKTTPVLDSDGSFLYSKL 436
QY 590 TVDKSRMGOQGVFSCSVMHAEALHNHYTQKSLSLSPG 625
DB 437 TVDKSRMGOQGVFSCSVMHAEALHNHYTQKSLSLSPG 472

RESULT 75
US-08-470-299-4
Sequence 4, Application US/08470299
Patent No. 5783181
GENERAL INFORMATION:
APPLICANT: Browne, Michael J.
APPLICANT: Murphy, Kay E.
APPLICANT: Chapman, Conrad G.
APPLICANT: Clinkenbeard, Helen E.
APPLICANT: Young, Peter R.
APPLICANT: Shatzman, Allan R.
TITLE OF INVENTION: No. 5783181el Compounds
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road, P.O. Box 1539
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```
APPLICATION NUMBER: US/08/470,299
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sutton, Jeffrey A.
REGISTRATION NUMBER: 34,028
REFERENCE/DOCKET NUMBER: P31005C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5024
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-470-299-4
```

```
Query Match 35.7%; Score 1220.5; DB 1; Length 387;
Best Local Similarity 83.4%; Pred. No. 6.8e-89;
Matches 236; Conservative 13; Mismatches 23; Indels 11; Gaps 4;
```

```
QY 351 VSKKEKPVWVLPNPKMQCLISDGGVLLSNIKVLPT--WSTFVPCPA-PEPKSCDKT 407
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 107 LKRLDRNLMLGL---AGLNSCPVKEANQSTLEFLERLKTIMREKDKSCSGTEPKSADKT 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 408 HTCC-----PELLGSPVFLFPKPKDITMISRTPEVTGVVDVSHEDPEVKNVYDQVE 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 164 HTCPPCPAPELLGGSPVFLFPKPKDITMISRTPEVTGVVDVSHEDPEVKNVYDQVE 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 463 VHNATKPREQYNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ 522
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 224 VHNATKPREQYNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 523 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTIPVLDSGS 582
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 284 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTIPVLDSGS 343
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 583 FFLYSKLTVDKSRMQGVFSCSVNHEALHNHYTQKSLSLSPG 625
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 344 FFLYSKLTVDKSRMQGVFSCSVNHEALHNHYTQKSLSLSPG 386
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 76

```
US-09-532-856-6
Sequence 6, Application US/09532856
Patent No. 6458350
GENERAL INFORMATION:
APPLICANT: COSMAN, David J.
APPLICANT: MULLBERG, Jurgen H.
APPLICANT: FANSLAW III, William C.
APPLICANT: KUBIN, Marek
TITLE OF INVENTION: ULBP DNA AND POLYPEPTIDES
FILE REFERENCE: 2866-USA
CURRENT APPLICATION NUMBER: US/09/532,856
PRIOR FILING DATE: 2000-03-22
EARLIER APPLICATION NUMBER: PCT/US98/27048
EARLIER FILING DATE: 1998-12-17
EARLIER APPLICATION NUMBER: 60/069,857
EARLIER FILING DATE: 1997-12-17
EARLIER APPLICATION NUMBER: 60/092,946
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 453
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)..(223)
OTHER INFORMATION: ULBP-2 sequences
FEATURE:
```

```
NAME/KEY: PEPTIDE
LOCATION: (224)..(453)
OTHER INFORMATION: Human Ig Fc sequences
US-09-532-856-6
```

```
Query Match 35.7%; Score 1220.5; DB 4; Length 453;
Best Local Similarity 58.9%; Pred. No. 8.6e-89;
Matches 258; Conservative 31; Mismatches 70; Indels 79; Gaps 11;
```

```
QY 236 GELWQAEKRRASSKSMITFDLAKNEVSKRTYDQPKLQMGKKLPLHLTPALPOYVAGSG 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 46 GPRWCAVQGVDEKTFPHYDGNKTVT-----FVSPLGKTL----- 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 NLTLAEAKTGKHOENVLVWRATQKVLTCGVWGPTSP-KLMLSLKLENKKAQVSK- 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 82 NVTIAMKAKNPVLEKREVDIL---TEQURDIOLENYTPKEPLTLQARNSCQKAEHSGG 137
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 -----REKPVW-VLNPEA-----GMWQCL--LSDS 375
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 138 SWQSPDQGIPLFDSEKRWMTTVHPGARKKKEKENDKVAMS FHYFSMGDCIGWLED- 196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 376 GQVLLSNIKVLPTWSTFVPCPADEP--KSCDKHTTC-----PELLGSPVFLFPKPK 427
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 197 --FLMGMDSTLEPESAGAPLAMS SGTTLRRSCDKTHTCPCPADAEAGAPSVFLFPKPK 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 428 DTLMISRTPEVTGVVDVSHEDPEVKNVYDQVEVHNATKPREQYNSTYRVSVLTV 487
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 255 DTLMISRTPEVTGVVDVSHEDPEVKNVYDQVEVHNATKPREQYNSTYRVSVLTV 314
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 488 LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGPREQVYTLPPSRDELTKNQVSLTCL 547
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 315 LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGPREQVYTLPPSRDELTKNQVSLTCL 374
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 548 VKGFYPSDIAVEMESNQGPENNYKTTIPVLDSGSFFLYSKLTVDKSRMQGVFSCSV 607
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 375 VKGFYPSDIAVEMESNQGPENNYKTTIPVLDSGSFFLYSKLTVDKSRMQGVFSCSV 434
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 608 HEALHNHYTQKSLSLSPG 625
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 435 HEALHNHYTQKSLSLSPG 452
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 77

```
US-09-524-100C-6
Sequence 6, Application US/09524100C
Patent No. 6653447
GENERAL INFORMATION:
APPLICANT: COSMAN, David J.
APPLICANT: MULLBERG, Jurgen H.
APPLICANT: FANSLAW III, William C.
APPLICANT: KUBIN, Marek
APPLICANT: ARMSTRONG, Richard J.
TITLE OF INVENTION: ULBP DNA AND POLYPEPTIDES
FILE REFERENCE: 2866-US
CURRENT APPLICATION NUMBER: US/09/524,100C
PRIOR FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: PCT/US98/27048
PRIOR FILING DATE: 1998-12-17
PRIOR APPLICATION NUMBER: US 60/069,857
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: US 60/092,946
PRIOR FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 453
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (1)..(223)
```

```

: OTHER INFORMATION: UBP-2 sequences
:
: FEATURE:
: NAME/KEY: PEPTIDE
: LOCATION: (224)..(453)
:
: OTHER INFORMATION: Human Ig Fc sequences
US-09-524-100C-6

```

Query Match	35.7%	Score 1220.5;	DB 4;	Length 453;
Best Local Similarity	-58.9%	Pred. No. 8.6e-89;		
Matches 258; Conservative	31;	Mismatches 70;	Indels 79;	Gaps 11;

Oy	236	GELMMOABERSSSSKMITEDLNKKEVSXKRVLPQDPKLOMGKKLPLHLTLPOLRYVAGSG	235
Db	46	GPRMCVAGQGVDEKTLIHYDCGNKIYT-----FVSPLGKTL-----	81
Oy	296	NLTFLAEAKTGKLDHENVLMVRATOLXNLTCEVWGPSP-KLMLSLKENKAVSK-	353
Db	82	NTTAMKAKQRPVLRVAVDIL-----TEQJLDIQLENVYTERPEBLTQARMSCKQKEGHSIG	137
Oy	354	-----REKPVW-VLNDEA-----GMNGCL--LSDS	375
Db	138	SKQSPSPDGOIFLLPDESKMMWTTTNGAKRKKKKENDKVVAMSGFNHFSMDCCJMWLED-	196
Oy	376	GOVLLSNIKIVLEPTWSTPVCBAPEB--KSCDKLTHTC-----DELGSPVFLPREPKX	427
Db	197	--FLMGMDSTLEPSACAPLAMSGGTQLRSCDKTHTCRPCAPABAGAPSVFLPREPKX	254
Oy	428	DTLMTISTPREVTVVVVDVSHEDPEYKFMNYDGVGVHNAKTKPREEOQNSTYRVVSLTV	487
Db	255	DTLMTISTPREVTVVVVDVSHEDPEYKFMNYDGVGVHNAKTKPREEOQNSTYRVVSVLTV	314
Oy	488	LHODMLNGKXYCKVCVSNKALPAPIEKTISKAGQREPOVYTLPPSRDELTKNOVSLTCL	547
Db	315	LHODMLNGKXYCKVCVSNKALPAPIEKTISKAGQREPOVYTLPPSRDELTKNOVSLTCL	374
Oy	548	VKGFYPSDIAVEMESNGQPENNYKTTPPVLYDSGSFFLYSKLTVDKSRWQGNVFCSCVM	607
Db	375	VKGFYPSDIAVEMESNGQPENNYKTTPPVLYDSGSFFLYSKLTVDKSRWQGNVFCSCVM	434
Oy	608	HEALHNHYTOKSLSLSPG	625
Db	435	HEALHNHYTOKSLSLSPG	452

```

RESULT 78
US-09-499-846-10
: Sequence 10, Application US/09499846
: Patent No. 6656728
: GENERAL INFORMATION:
: APPLICANT: Kanaugh et al.
: TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
: FILE OF INVENTION: RECEPTOR-IMMUNOGLOBULIN FUSION
: FILE REFERENCE: 035784/195012 (5784-
: CURRENT APPLICATION NUMBER: US/09/499,846
: CURRENT FILING DATE: 2000-02-07
: NUMBER OF SEQ. ID NOS: 12
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 10
: LENGTH: 497
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-499-846-10

```

```

Query Match Similarity      35.7% ; Score 1220.5 ; DB 4 ; Length 497 ;
Best Local Similarity     65.7% ; Pred No. 9.8e-89 ;
Matches    253 ; Conservative   18 ; Mismatches 56 ; Indels   56 ; Gaps    8

QY          293 GSGNLTLLAEAKTGKLDH--EVNVVVMKATOLGNL--TCEWGPSTSKMLSIKLE-   345
            |||         |||||        |::|       :||:
DB         118 GSINHTQLDVERSPHRPILOAGLRPAKTVALGSNVEMCKVSDDPGCHIQWLKHIEVN   177
            |||         |||||        |::|       :||:

QY          346 -----NKEAKVSKREKPVVNL-----PEAGMOCCLSDSGGVLLLESNIK   385

```

```

Db      178 GSKIGDNLPRYYQILKTAGVNTDDKEMVLIHLRVNFEFDAGEEYCLAGNISIGLSHS--- 234
Qy      386 VLPTWT-----PVCPAP-----EPKSCDKTHHC-----PELLGGSVF 420
                | |
Db      235 ---AMLTIVLEAEERPAVMTSEFLYIEGSGSGLOEPKSCDKTHHCPCPAPALEGGSVF 291
Qy      421 LPPPKKDTLMTISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAATKPREQYNSTYR 480
Db      292 LPPPKKDTLMTISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAATKPREQYNSTYR 351
Qy      481 VNSVLTVLHODMLNGEKYKCKVSNALPAPLEKITSAKGQPREPQVYTLPPSDELTKN 540
Db      352 VNSVLTVLHODMLNGEKYKCKVSNALPAPLEKITSAKGQPREPQVYTLPPSDELTKN 411
Qy      541 QVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFPLYSLKLVYDSKRWQGN 600
Db      412 QVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFPLYSLKLVYDSKRWQGN 471
Qy      601 VFSCSVMEHIALNHNTQKSLSPG 625
Db      472 VFSCSVMEHIALNHNTQKSLSPG 496

```

```

RESULT 79
US-09-499-846-8
: Sequence 8, Application US/09499846
: Patent No. 6656728
: GENERAL INFORMATION:
: APPLICANT: Kavanaugh et al.
: TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
: FILE REFERENCE: RECEPTOR-IMMUNOGLOBULIN FUSION
: CURRENT APPLICATION NUMBER: US/09/499,846
: CURRENT FILING DATE: 2000-02-07
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 8
: LENGTH: 525
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-499-846-8

```

	Query Match	35.7%;	Score 1220.5;	DB 4;	Length 525;
	Best Local Similarity	65.7%;	Pred. No. 1.1e-88;		
	Matches 253;	Conservative 18;	Mismatches 56;	Indels 58;	Gaps 8
QY	293	GGSNLTALAEAKTKLHQ--EVLNVVRATOLQKLT--TCVWGPTSPKLMLSIKLF-	345		
Db	146	GSINHTYQLDIVERSPHRPILQACLIPANKTYALGNSNVEFMCKVYSDPQHILQKHIVN	205		
QY	346	-----NKEAKYSKRKKRPVVNL-----PEAGMOCILSDSQVLLSENIK	385		
Db	206	GSKTGPDNLPYVQILKTRAGVTTDKEMEVHLNRVNSFEDAEHYTCLAGNSIGLSHNS--	262		
QY	386	VLPTMTST-----PYVCPAP-----EPKSCDXTHTC-----DELLGSPSVF	420		
Db	263	---AMLTVLALBERPAPMTSPFLYLEGSGPGLQEPKSCDKTHTCPCPAPAELEGSPSVF	319		
QY	421	LFPPKPKDMLISTRPETVCVVNVDSHDEPVKKNMYVDGVEVNNAKTKPRBEQVNSTYR	480		
Db	320	LFPPKPKDMLISTRPETVCVVNVDSHDEPVKKNMYVDGVEVNNAKTKPRBEQVNSTYR	379		
QY	481	VSVLTVLTHQMLNGKEYCKCVSNKALPAPIEKTISKAKQPREPOVYTLTPPSRDELTKN	540		
Db	380	VSVLTVLTHQMLNGKEYCKCVSNKALPAPIEKTISKAKQPREPOVYTLTPPSRDELTKN	439		
QY	541	QVSLTCLVKGFPESDIAVEMESNQPENNYKTPPVLDDSGSFLYSKLTVDKSKWQGN	600		
Db	440	QVSLTCLVKGFPESDIAVEMESNQPENNYKTPPVLDDSGSFLYSKLTVDKSKWQGN	499		
QY	601	VFSCSVMEALHNHYTKSLSLSPG	625		
Db	500	VFSCSVMEALHNHYTKSLSLSPG	524		

```
RESULT 80
US-09-131-247-16
; Sequence 16, Application US/09131247
; Patent No. 6294170
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C.
; APPLICANT: Herhenson, Susan
; APPLICANT: Bevilacqua, Michael P.
; APPLICANT: Collins, David S.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
; FILE REFERENCE: A-365F
; CURRENT APPLICATION NUMBER: US/09/131,247
; EARLIER FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: 60/055,185
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: PCT/US 97/02131
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Human
US-09-131-247-16
```

```
Query Match      35.7%; Score 1220; DB 3; Length 388;
Best Local Similarity 65.9%; Pred. No. 7.5e-89;
Matches 255; Conservative 23; Mismatches 73; Indels 36; Gaps 8;
```

```
QY 253 TFDKKNKVSXRYTODPKLQMGKLLPLHLTPQALPOYAGSGNLTALAEKTKLHQEV 312
DB 23 TFIYRNQQL-VAGYLQGFNVNLEBKIDVPIEPHALFLGHGKMLSC-VKSGD-ETRL 79
QY 313 NLVYMRATQLOKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPVWNLNPEA--GMMQC 370
DB 80 QLEAVNTIDISEN-----RKQDKRFPFIRSDSGPTTSFESAACPGWFLC 123
QY 371 LLSSGQVLESNTIKVLTFTWSTPVPC-----PAPEPKSCDKHTTC-----PELLGSPS 418
DB 124 TAMADQPVSLTN--MEDEGVMTKFFQDEDAAEKPSDKHTTCCPPCAPPELLGSPS 180
QY 419 VFLEPPKPKDQTLMTSRPEVTCVAVDVSHDEPEVKFMYVGVVNAKTKPREEQNST 478
DB 181 VFLEPPKPKDQTLMTSRPEVTCVAVDVSHDEPEVKFMYVGVVNAKTKPREEQNST 240
QY 479 YRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDEL 538
DB 241 YRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDEL 300
QY 539 KNOVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMQ 598
DB 301 KNOVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMQ 360
QY 599 GNVFSCVMEHALNHYTKSLSPG 625
DB 361 GNVFSCVMEHALNHYTKSLSPG 387
```

```
RESULT 81
US-09-101-593-18
; Sequence 18, Application US/09301593A
; Patent No. 6455677
; GENERAL INFORMATION:
; APPLICANT: Park, John E.
; APPLICANT: Garin-Chees, Pilar
; APPLICANT: Bamberger, Uwe
; APPLICANT: Legier, Olivier
; APPLICANT: Saldana, Jose W.
; APPLICANT: Reutig, Wolfgang J.
; TITLE OF INVENTION: FAP-specific Antibody with Improved Productibility
```

```
; FILE REFERENCE: 0652.1890001
; CURRENT APPLICATION NUMBER: US/09/301,593A
; CURRENT FILING DATE: 1999-04-29
; EARLIER APPLICATION NUMBER: EP 98107925.4
; EARLIER FILING DATE: 1998-04-30
; EARLIER APPLICATION NUMBER: US 60/086,049
; EARLIER FILING DATE: 1998-05-18
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-101-593-18
```

```
Query Match      35.7%; Score 1219.5; DB 4; Length 453;
Best Local Similarity 46.1%; Pred. No. 1e-88;
Matches 281; Conservative 42; Mismatches 107; Indels 179; Gaps 16;
```

```
QY 30 LGKKDVTVELTCTASQKSIQF--HMKNSQIKILNQSGF-LTKGPSKLNDRADSRSL 86
DB 10 LVKPGASVKNQSCKTSRTTFETITIMVYQSHGSLIEWIGINPNNGIIPNTNOKFKGATL 69
QY 87 W---DQGNFPLIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSPTHLIQ--GQSLTL 141
DB 70 TVGKSSSTAYMELRSLTSEDSAVYFC-----ARRIAVGY--DEGHAMDYWGQGTSTV 119
QY 142 TLSPSSGSSPVQCRSPRGKNIQCGKTLVSQLELDGSGTWTCTVLONQKKEFKIDIV 201
DB 120 TVSSASTGKSVPLAPSSKSTSGG--TALGCL----- 151
QY 202 LAFQKASSIYKKGEVEFSFPLAFVVELTSGELMQAERASSKSWITDKNKEV 261
DB 152 -----VKDYFPEPVTVS-----MNSGALTSG----- 172
QY 262 SVKRVITODPKLQMGKLLPLHLTPQALPOYAGSGNLTALAEKTKLHQEVNLVYMRATQ 321
DB 173 -----VH-TTPAVL-QSSGLYSLSSVTVPSSSLGQTYYI----- 205
QY 322 LQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPVWNLNPEAGMQLLSSGQVLE 381
DB 206 -----CNV-----NHKPSMTKYDKV----- 221
QY 382 SNIKVLPYTWSTPVPCAPPEPKSCDKHTTC-----PELLGSPVFLPPPKDQTLMTSRTP 436
DB 222 -----EPKSCDKHTTCCPPCAPPELLGSPVFLPPPKDQTLMTSRTP 263
QY 437 EYTCVAVDVSHDEPEVKFMYVGVVNAKTKPREEQNSTYRVVSVLTVLHQDWLNGK 496
DB 264 EYTCVAVDVSHDEPEVKFMYVGVVNAKTKPREEQNSTYRVVSVLTVLHQDWLNGK 323
QY 497 EYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDI 556
DB 324 EYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDEMTKNOVSLTCLVKGFPYSDI 383
QY 557 AVEMESNGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMQGNVSCSMHEALNHYT 616
DB 384 AVEMESNGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMQGNVSCSMHEALNHYT 443
QY 617 QKSLSLSPG 625
DB 444 QKSLSLSPG 452
```

```
RESULT 82
US-09-740-002-27
; Sequence 27, Application US/09740002
; Patent No. 6537809
; GENERAL INFORMATION:
; APPLICANT: BRAMS, PETER
; APPLICANT: MORROW, PHILLIP
; TITLE OF INVENTION: NEUTRALIZING HIGH AFFINITY HUMAN MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: SPECIFIC TO RSV F-PROTEIN AND METHODS FOR THEIR
```

```

: TITLE OF INVENTION: MANUFACTURE AND THERAPEUTIC USE THEREOF
: FILE REFERENCE: 037003-0275759
: CURRENT APPLICATION NUMBER: US/09/740,002
: CURRENT FILING DATE: 2000-12-20
: PRIOR APPLICATION NUMBER: 09/335,697
: PRIOR FILING DATE: 1999-06-18
: PRIOR APPLICATION NUMBER: 08/488,376
: PRIOR FILING DATE: 1995-06-07
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 27
: LENGTH: 475
: TYPE: PR1
: ORGANISM: Homo sapiens
US-09-740-002-27

Query Match      35.7%; Score 1219.5; DB 4; Length 475;
Best Local Similarity 45.0%; Pred.No.1,1e-88;
Matches 290; Conservative 31; Mismatches 116; Indels 207; Gaps 16

QY      10 LLLVQLALLPAATGKRVYLGGKADYELTCTASQKKSICPHMKNNOIKLNGSFL 69
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      10 LVAATRVLSQVQLQESGPALVKPQTLLTCTFS-----GFSLSTRGMSVMI 58
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      70 TKGSKLND---RDSRRSLMDQGNF-----PLIKLKIEDSDT 106
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      59 RQPKALEMLARID---WDDTFYASLKTSLISKDTISKDTSKNQVLRMTVDDVDTAT 113
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      107 YICEVEDQKEEYQLVFLGLTANSDDLHQQGSLTLLTLESPSSPSVOCRSPRGHIOG 166
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      114 YFCARASLYDSDSFLF-----YHAWGCGTIVTVSASATKGSGSVFPLPASSKSTSG 166
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      167 KTLVSQLELDGSGTWCTVTLQNKQKVEPKIDVLLAFQKASSIVYKKEGEOVEFSPLA 226
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      167 -TAALGCL-----YKDYFPEP 181
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      227 FTVEKLTGSGELMWQAERASSKSWITFDLKNKEVSVKRVYQDPKLQMGKKLPLHLTLPLQ 286
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      182 VTVS-----NMSGALITSG-----VH-TTPA 200
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      287 ALPQVAGSGNLTLLAEATKGKLIHQEVLVVMRATQLOKNLTGCEVMGPTSPKMLSLKLEN 346
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      201 VL-QSSGLXLSVTVTVSSSLGTQTYI-----CNV----- 230
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      347 KEAKVSKREKPYVWLVNPEAGMWQCLLSDGQVLLSNIKVLPTWSTPVPCEAPEKSCDK 406
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      231 -----NHKP-----SNRKV-----DKAEPEKSCDK 250
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      407 THTC-----PELGGPSVFLPPPKRKDTLMTSRTEPVTCVVVDVSHEDPEVFENNYVGV 461
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      251 THTCPCPAPELLGGPSVFLPPEPKDKDTLMTSRTEPVTCVVVDVSHEDPEVFENNYVGV 310
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      462 EVHNAKTPREEOVNSTYRVVSVLVTLVHODMLNGEKYCKVSNKALPAPIETISKAKQ 521
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      311 EVHNAKTPREEOVNSTYRVVSVLVTLVHODMLNGEKYCKVSNKALPAPIETISKAKQ 370
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      522 PREPOVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVEMESGOGENNYKTTIPVLDSDG 581
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      371 PREPOVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVEMESGOGENNYKTTIPVLDSDG 430
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      582 SFFLYSKLTVDKSRWQOGNVFSCSYMHREALHNHYTKSLSPG 625
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      431 SFFLYSKLTVDKSRWQOGNVFSCSYMHREALHNHYTKSLSPG 474
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 83
US-08-487-550-12
: Sequence 12, Application US/08487550
: Patent No. 6113898
: GENERAL INFORMATION:
: APPLICANT: Anderson, Dattaji R.
: TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
: TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,"

```

```

1  TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF
2  TITLE OF INVENTION: IMMUNOSUPPRESSANTS"
3  NUMBER OF SEQUENCES: 12
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
6  STREET: 699 Prince Street
7  CITY: Alexandria
8  STATE: VA
9  COUNTRY: USA
10 ZIP: 22314
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC-DOS/MS-DOS
15 SOFTWARE: Patent In Release #1.0, Version #1.30
16 CURRENT APPLICATION DATA:
17 FILING DATE: 07-JUN-1995
18 CLASSIFICATION: 435
19 ATTORNEY/AGENT INFORMATION:
20 NAME: Teskin, Robin L.
21 REGISTRATION NUMBER: 35,030
22 REFERENCE/DOCKET NUMBER: 012712-131
23 TELECOMMUNICATION INFORMATION:
24 TELEPHONE: 703-836-6620
25 TELEFAX: 703-836-2021
26 INFORMATION FOR SEQ. ID NO: 12:
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 476 amino acids
29 TYPE: amino acid
30 TOPOLOGY: linear
31 MOLECULE TYPE: protein
32 US-08-487-550-12
33
34 Query Match 35.7%; Score 1219.5; DB 3; Length 476;
35 Best Local Similarity 53.7%; Pred. No. 1.1e-88;
36 Matches 277; Conservative 31; Mismatches 89; Indels 119; Gaps 15;
37
38 Oy 171 VSGLEIQQDSG-----TWTCYVLQNKQKVEREKIDIVLAFQKASIVYKKEGQVE 220
39 Db 18 LSOVOQAESQPGIWKPSFTLSLTCV-----SGSGSIS 49
40 Oy 221 FSPPLAFYVEKLTGSGELMMQAEBSASSKSWITFDLKNKEVSVKRVTQDPKLQMGKKPL 280
41 Db 50 GGGWGCM-IRQPGKLENIIGSFYSSGNTYVNSPLKS-QVITIS--TDTSKNQPSLKL-- 103
42 Oy 281 HLTLPQALPOLYAGSGVLTALAE---KTKGLHGEVNLVWKEATOLQKULTCEWGP--- 333
43 Db 104 -----NNTAADTAIVYVCVADRLEFSVGMVY-----NNMFDVWGPGVL 141
44 Oy 334 -----TSPKMLSLKLENKEAKVSKR-----EKPVWV-----LNPEAGMOC 370
45 Db 142 VTVSASASTKQPSVFPIAPSSKSTSGGTALGCLVKDYPPEPEPTVWNSGALTSGVHTPEA 201
46 Oy 371 LLSDSQVLLS-----NIXKLPTMSTFVPCPABEPKSCDTHNC----- 410
47 Db 202 VLOSSGGLSYSSVVTVPSSSLGTQYICVNHKPS-NTVVDKKA-EPKSCDTHNCPPCP 259
48 Oy 411 -PELLGSPSVFLPPPKDQTLMSRPEVTCVVDVSHEDPEVKFNWYVDGEVNAKTK 469
49 Db 260 APELLGSPSVFLPPPKDQTLMSRPEVTCVVDVSHEDPEVKFNWYVDGEVNAKTK 319
50 Oy 470 PREQYNSTYRVVSVLTVLHQDLNMEKEKCVSKNALPAPIEKTSIAKQGPREFQVYT 529
51 Db 320 PREQYNSTYRVVSVLTVLHQDLNMEKEKCVSKNALPAPIEKTSIAKQGPREFQVYT 379
52 Oy 530 LPPSRDELTKNQVSLTCLVKGFYPSDIAVEGSSNGQPENNYKTTTPPVLDSDGSFLYSK 589
53 Db 380 LPPSRDELTKNQVSLTCLVKGFYPSDIAVEGSSNGQPENNYKTTTPPVLDSDGSFLYSK 439
54 Oy 590 TVDKSRMOQGNVFSGVMEHALLHNHTQKSLISLSPG 625
55 Db 440 TVDKSRMOQGNVFSGVMEHALLHNHTQKSLISLSPG 475

```


RESULT 84

US-09-526-098-12

Sequence 12, Application US/09526098

Patent No. 6492134

GENERAL INFORMATION:

APPLICANT: Anderson, Darrell R.

TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF."

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS

TITLE OF INVENTION: IMMUNOSUPPRESSANTS"

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: 699 Prince Street

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/526,098

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/383,916

FILING DATE:

APPLICATION NUMBER: US 08/487,550

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 012712-131

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6620

TELEFAX: 703-836-2021

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 476 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-526-098-12

Query Match 35.7%; Score 1219.5; DB 4; Length 476;

Best Local Similarity 53.7%; Pred. No. 1.1e-88;

Matches 277; Conservative 31; Mismatches 89; Indels 119; Gaps 15;

171 VSQELQDSG-----TWCTVLQNKVKFEKIDIVLAFQKASSIYKKEGEQVE 220

18 LSQVQLQSGGGLVPSFETLSLTCAV-----SGGSIS 49

221 FSPFLATVEKLTGSGELMWAERASSKSWITPDLNKKEYSVKRVTDPLQNGKLLPL 280

50 GGYGNGW-IRQPPGGLGEMISSFYSSSGNTYNSPLKS-QVTIS--TDTSKQFSKL- 103

281 HLTLPQALPOYAGSGNLTFLAE---KTGKLHQEVNLYVMARATQQLKULCEVWGP- 333

104 -----NSMTAADTAVYYCVRDLFSVGVGVY-----INMFVWAGPEVL 141

334 -----TSPKMLSLKLENKEAKVSKR-----EKPVVY-----LNPEAGMWQC 370

142 VTSSASTKGPSVFLAPSSKSTSGTAAGLVKDYPPPEVYTSWMSGALTSVHPFA 201

371 LLSBSGQVLES-----NIKVLPTWSTFVPCPAPPEKSCDKHTHC--- 410

202 VLQSSGLVSLSSVYTVTPSSSLGTOTYICNVNPKS-NTKVDKKA-EPKSCDKHTCPSP 259

Qy 411 -PELLGSPVFLPEPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTK 469

Db 260 APELLGSPVFLPEPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTK 319

Qy 470 PREQYNSTRVSVLVTLHODWLNKGEYKCKVSNKALPAPIETISKAKQPREPOVYT 529

Db 320 PREQYNSTRVSVLVTLHODWLNKGEYKCKVSNKALPAPIETISKAKQPREPOVYT 379

Qy 530 LPPRDELTKNQVSLTCLVGFPSDIAVEMESNGQPENNYKTPPVLDSDGSFFLYSKL 589

Db 380 LPPRDELTKNQVSLTCLVGFPSDIAVEMESNGQPENNYKTPPVLDSDGSFFLYSKL 439

Qy 590 TVDKSRWQGVNPFSCSYVMEALNHNHYTQKSLSPG 625

Db 440 TVDKSRWQGVNPFSCSYVMEALNHNHYTQKSLSPG 475

RESULT 85

US-08-472-888A-7

Sequence 7, Application US/08472888A

Patent No. 6613746

GENERAL INFORMATION:

APPLICANT: Seed, Brian

APPLICANT: Walz, Gerd

TITLE OF INVENTION: AGP-ANTIBODY FUSION PROTEINS

TITLE OF INVENTION: AND RELATED MOLECULES AND METHODS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark & Eibing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Pastesd for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,888A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/618,314

FILING DATE: 23-NOV-1990

ATTORNEY/AGENT INFORMATION:

NAME: Eibing, Karen L.

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/258001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 442 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-472-888A-7

Query Match 35.6%; Score 1216; DB 4; Length 442;

Best Local Similarity 89.6%; Pred. No. 1.9e-88;

Matches 233; Conservative 3; Mismatches 14; Indels 10; Gaps 2;

371 LLSBSGQVLESNIKVLPTWSTFVPCPAPPEKSCDKHTHC-----PELLGSPVFLPEPK 425

Db 187 VLQSSGLVSLSSVYTVTPSSSDKRV-----EPKSCDKHTHTPCPAPPELLGSPVFLPEPK 241

Qy 426 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKQPREQYNSTRVSVLV 485

Db 242 PKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYVDGVEVHNAKTKRPREQNSTYRVVSVL 301
QY 486 TVLHODMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPQVYTLTPSRDELTKNQVSLT 545
Db 302 TVLHODMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPQVYTLTPSRDELTKNQVSLT 361
QY 546 CLVKGFPYSDIAVESMSGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMOQGNVFSGS 605
Db 362 CLVKGFPYSDIAVESMSGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMOQGNVFSGS 421
QY 606 VMHEALHNYTKSKLSLSPG 625
Db 422 VMHEALHNYTKSKLSLSPG 441

RESULT 86
PCT-US96-10043-9
Sequence 9, Application PC/TUS9610043
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES
TITLE OF INVENTION: AND METHODS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10043
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,213
FILING DATE: 14-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lech, Karen F.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 00786/284001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 442 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-10043-9

Query Match 35.6%; Score 1216; DB 5; Length 442;
Best Local Similarity 89.6%; Pred. No. 1.9e-88;
Matches 233; Conservative 3; Mismatches 14; Indels 10; Gaps 2;

QY 371 LLSDSGQVLLSNNKVLPTWSTPVPCEAPPEKSCDKHTHC-----PELLGSPSVLFPKP 425
Db 187 VLQSSGSLYSLSVTVVSSSDKKV-----EPKSCDKHTCPCEAPPELLGSPSVLFPKP 241
QY 426 PKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYVDGVEVHNAKTKRPREQNSTYRVVSVL 485
Db 242 PKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYVDGVEVHNAKTKRPREQNSTYRVVSVL 301
QY 486 TVLHODMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPQVYTLTPSRDELTKNQVSLT 545

Db 302 TVLHODMLNGKEKCKVSNKALPAPIEKTISKAKGQPREPQVYTLTPSRDELTKNQVSLT 361
QY 546 CLVKGFPYSDIAVESMSGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMOQGNVFSGS 605
Db 362 CLVKGFPYSDIAVESMSGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMOQGNVFSGS 421
QY 606 VMHEALHNYTKSKLSLSPG 625
Db 422 VMHEALHNYTKSKLSLSPG 441

RESULT 87
PCT-US96-10043-11
Sequence 11, Application PC/TUS9610043
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES
TITLE OF INVENTION: AND METHODS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10043
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,213
FILING DATE: 14-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lech, Karen F.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 00786/284001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-10043-11

Query Match 35.6%; Score 1215.5; DB 5; Length 437;
Best Local Similarity 97.0%; Pred. No. 2e-88;
Matches 227; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 397 PAPEPKSCDKHTHC-----PELLGSPSVLFPKPCKDTLMISRTPEVTCVVVDVSHEDPE 451
Db 203 PEGEPKSCDKHTHCPCPAPELLGSPSVLFPKPCKDTLMISRTPEVTCVVVDVSHEDPE 262
QY 452 VRFNMWYVDGVEVHNAKTKRREQNSTYRVVSVLTVLHODMLNGKEKCKVSNKALPAPI 511
Db 263 VRFNMWYVDGVEVHNAKTKRREQNSTYRVVSVLTVLHODMLNGKEKCKVSNKALPAPI 322
QY 512 EKTISKAKGQPREPQVYTLTPSRDELTKNQVSLTCLVKGFPYSDIAVESMSGQPENNYK 571
Db 323 EKTISKAKGQPREPQVYTLTPSRDELTKNQVSLTCLVKGFPYSDIAVESMSGQPENNYK 382

Qy 572 TTPVLDSGSEFLYSKLTVDKSRMOQGNVSCSVMEHALHNHYTKSLSPG 625
Db 383 TTPVLDSGSEFLYSKLTVDKSRMOQGNVSCSVMEHALHNHYTKSLSPG 436

RESULT 88
US-08-378-939-10
; Sequence 10, Application US/08378939
; Patent No. 5876961
; GENERAL INFORMATION:
; APPLICANT: CROME, JAMES SCOTT
; APPLICANT: LEWIS, ALAN PETER
; TITLE OF INVENTION: PRODUCTION OF ANTIBODIES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KIRZ
; STREET: 555 THIRTEENTH ST. N.W.
; CITY: WASHINGTON
; STATE: D. C.
; COUNTRY: U.S.
; ZIP: 20004

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,939
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/952640
; FILING DATE: 01-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ERNST, BARBARA G
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1808-118
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 783-6040
; TELEFAX: (202) 783-6031
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-939-10

Query Match 35.6%; Score 1215; DB 2; Length 476;
Best Local Similarity 97.8%; Pred. No. 2.5e-88;
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

Qy 400 EPKSCDKHTHC-----PELLGSPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEKXF 454
Db 245 EPKSCDKHTHCPCPAPELLGSPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEKXF 304
Qy 455 NMVYDGVVNAKATPREEOYNSTRVVSUTLVTHQDMLNGEKYCKSNKALPAPIEKT 514
Db 305 NMVYDGVVNAKATPREEOYNSTRVVSUTLVTHQDMLNGEKYCKSNKALPAPIEKT 364
Qy 515 ISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVGFYPSDIAVEMESNGQPENNYKTTP 574
Db 365 ISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVGFYPSDIAVEMESNGQPENNYKTTP 424
Qy 575 PVLDSGSEFLYSKLTVDKSRMOQGNVSCSVMEHALHNHYTKSLSPG 625
Db 425 PVLDSGSEFLYSKLTVDKSRMOQGNVSCSVMEHALHNHYTKSLSPG 475

RESULT 89
US-09-157-452B-12
; Sequence 12, Application US/09157452B
; Patent No. 6482409

GENERAL INFORMATION:
; APPLICANT: Lobb, Roy R.
; APPLICANT: Burkly, Linda C.
; TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 10274-004003
; CURRENT APPLICATION NUMBER: US/09/157,452B
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: US 08/950,660
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: US 08/373,857
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/284,603
; PRIOR FILING DATE: 1994-08-11
; PRIOR APPLICATION NUMBER: PCT/US93/00924
; PRIOR FILING DATE: 1993-02-02
; PRIOR APPLICATION NUMBER: US 07/835,139
; PRIOR FILING DATE: 1992-02-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-157-452B-12

Query Match 35.5%; Score 1213.5; DB 4; Length 446;
Best Local Similarity 52.4%; Pred. No. 3e-88;
Matches 270; Conservative 40; Mismatches 92; Indels 113; Gaps 14;

Qy 126 TANSDFHLQ--GOSTLTLESPPGSSP-----SVQCRSP-RGKNIOGG--KTLVSQLELQ 177
Db 29 TTPSRVLAQGDGVSLSCTGCGESPFPMRQIDBPLNGKVTNEGTSTLTNNPVSFG 88
Qy 178 DSGTWCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGQVE-FSPPLAFTVEKLTGSG 236
Db 89 NEHSYLTATCESRKLE-----KGIOVEIYSFP----- 116
Qy 237 ELWQARASSSKSWITFDLKNKEVSXKRYTQDPKLOMGKKLPLHLTLPOLPYAGSGN 296
Db 117 -----KDPRIHL-----SGPLAGKPIYKCSVADYRP----- 144
Qy 297 LTLLEAKTGKLEHVEVLMVVRATOLQKNLTCEVWGPTSPKMLSLKLENKAVSKREK 356
Db 145 -----FRLRLIDLLKGDHLMKSG--FELBEDARK-----SLFTKSLLEV----- 181
Qy 357 PVMVLNPEAGMOCLLSDSQVLL--ESNIVLPTWSTPVCPCPAPKSCDKHTHC----- 410
Db 182 -----FTFVIEDIGKVLVCRAKLHIDEMDSVPTVROAVKELQVDKHTCCPCPA 230
Qy 411 PELLGSPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEKKNWYDGVVNAKATKP 470
Db 231 PELLGSPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEKKNWYDGVVNAKATKP 290
Qy 471 REEOYNSTRVVSUTLVTHQDMLNGEKYCKSNKALPAPIEKTISKAKGQPREPOVYTL 530
Db 291 REEOYNSTRVVSUTLVTHQDMLNGEKYCKSNKALPAPIEKTISKAKGQPREPOVYTL 350
Qy 531 PPSRDELTKNOVSLTCLVGFYPSDIAVEMESNGQPENNYKTTPVPLDSGSEFLYSKLT 590
Db 351 PPSRDELTKNOVSLTCLVGFYPSDIAVEMESNGQPENNYKTTPVPLDSGSEFLYSKLT 410
Qy 591 VDKSRMOQGNVSCSVMEHALHNHYTKSLSPG 625
Db 411 VDKSRMOQGNVSCSVMEHALHNHYTKSLSPG 445

RESULT 90
US-08-458-516-13
; Sequence 13, Application US/08458516
; Patent No. 5777085
; GENERAL INFORMATION:
; APPLICANT: Co, Man Sung
; APPLICANT: Tso, J. Yun

```

; TITLE OF INVENTION: Humanized Antibodies Reactive with
; TITLE OF INVENTION: GPIIb/IIIa
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William M. Smith
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,516
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/059,159
; FILING DATE: 03-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-37-3
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-458-516-13

Query Match      35.5%; Score 1213; DB 1; Length 449;
Best Local Similarity 97.8%; Pred. No. 3.3e-88;
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 400 EPKSCDKTHTC-----PELGGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKF 454
DB 218 EPPSCDKTHTCPPCPAPPELLGGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKF 277
QY 455 NMVYDGEVHNAKTPREBOYNSTYRVSVLYLVHODMNGKEYCKVSNKALPAPIEKT 514
DB 278 NMVYDGEVHNAKTPREBOYNSTYRVSVLYLVHODMNGKEYCKVSNKALPAPIEKT 337
QY 515 ISRAKQPREPOVYTLPPSDELTKNOVSLTCLVKGFYPSDIAVEMESNQPPENNYKTPP 574
DB 338 ISRAKQPREPOVYTLPPSDELTKNOVSLTCLVKGFYPSDIAVEMESNQPPENNYKTPP 397
QY 575 PVLDSDSFLLYSKLTVDKSRMOQGNVFCSVHMEALHNHYTQKSLSLSPG 625
DB 398 PVLDSDSFLLYSKLTVDKSRMOQGNVFCSVHMEALHNHYTQKSLSLSPG 448

RESULT 91
US-09-485-737B-67
; Sequence 67, Application US/09485737B
; Patent No. 6350860
; GENERAL INFORMATION:
; APPLICANT: Buysse, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC SHOCK.
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015
; CURRENT APPLICATION NUMBER: US/09/485,737B
; CURRENT FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
```

```

; PRIOR APPLICATION NUMBER: EPO 98870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: Patent version 3.0
; SEQ ID NO 67
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC
US-09-485-737B-67

Query Match      35.5%; Score 1212.5; DB 4; Length 468;
Best Local Similarity 45.3%; Pred. No. 3.9e-88;
Matches 287; Conservative 42; Mismatches 102; Indels 203; Gaps 17;

QY 11 LVLQALPAPATQGNKVLGKGPVETLTASQKKSIOFHKNSNQIKLQNGSFLT 70
DB 17 VILSQVQLVQSGSE-----LKKPGASVKISCKAS--GTFITDYGKMWYKQARGQ---L 65
QY 71 KQPSKLNDRADSRRLMD--QGNFP-----LIINKLIEDSDYICEVEDQKEV 118
DB 66 KMWGMINTYGTGSTYVDPKGRFVPSLDTVSAAVLIQISLKAEDTATYFC----- 116
QY 119 QLLVFGLTANSDLHLQ--GQSLTLTLSPSSSPVQCRSPRGKNIQGGKTLTSVQL 176
DB 117 -----ARRGFYADYWGQGTIVVSSASTKGPSPVPLAPSSKSTSG--TVALGCL-- 165
QY 177 QDSGTWTCTVLQNKQKVEFKIDIVLAFQKASSIYKKEGQVEFSPLAFTVEKLTGSG 236
DB 166 -----VKKQYFPEPVYS----- 177
QY 237 ELWMQKERASSSKSWITFDLKNKEVSVKRVTPDKLQMGKQLPLHLTPQALPOYAGSN 296
DB 178 ---WNSGALTSG-----VH--TPPAVL--QSSGLYS 201
QY 297 LTLALEAKTKGLHGEVNLVVMRATOLQKRLTCEWGPSTPKMLSLKENKEKSKRER 356
DB 202 LSSVYVTPSSSLGTQYI-----CNV-----NHRKSNTKVKV 233
QY 357 PWWNLNPEAGMMQCLLSDSGQVLBSNLIKVLPTWSTPVPAPPEPKCDKTHTC-----P 411
DB 234 RV-----EPKSCDKTHTCPPCPAP 252
QY 412 ELGGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAKTKPR 471
DB 253 ELGGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAKTKPR 312
QY 472 EEOYNSTYRVSVLYLVHODMNGKEYCKVSNKALPAPIEKTISAKQPREPOVYTLPP 531
DB 313 EEOYNSTYRVSVLYLVHODMNGKEYCKVSNKALPAPIEKTISAKQPREPOVYTLPP 372
QY 532 PSRDELTKNOVSLTCLVKGFYPSDIAVEMESNQPPENNYKTPPVLDSDGSFLLYSKLT 591
DB 373 PSREMTKNOVSLTCLVKGFYPSDIAVEMESNQPPENNYKTPPVLDSDGSFLLYSKLT 432
QY 592 DKSRMOQGNVFCSVHMEALHNHYTQKSLSLSPG 625
DB 433 DKSRMOQGNVFCSVHMEALHNHYTQKSLSLSPG 466

RESULT 92
US-09-485-737B-90
; Sequence 90, Application US/09485737B
; Patent No. 6350860
; GENERAL INFORMATION:
; APPLICANT: Buysse, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC SHOCK.
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015
```

```
; CURRENT APPLICATION NUMBER: US/09/485, 737B
; CURRENT FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 96/05165
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: EP0 96870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EP0 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC
US-09-485-737B-90
```

```
Query Match 35.5%; Score 1212.5; DB 4; Length 711;
Best Local Similarity 45.3%; Pred. No. 7.1e-88;
Matches 287; Conservative 42; Mismatches 102; Indels 203; Gaps 17;
```

```
QY 11 LVLVQLALLPRAATGKNNVLLGKGDVLELTCTASQKSIQFMKNSQIKILNQSGFLLT 70
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
17 VILSQVOLVWGSGE-----LKKPGASVXISCKAS---GYFTDYGMWVVKQAPGQ---L 65

QY 71 KGPBKLDNRADRSRLMD--QGNFP-----LIIKNLKITDSDTYICEVEDQKEV 118
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
66 KMMGMINTYTGESTYVDKGRFVFSLDTSVAAYLIQSLKADDTATYFC----- 116

QY 119 QLVFGLTANSDTLIQ--GQSLTLTLESPPGSSPVOCRRPGKNIQGGKTLISVSOLEL 176
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 -----ARGFAMDYWGCGTIVTVSSASTKGPSVFPPLAPSSKISG--TALAGCL-- 165

QY 177 QDSGTWTCTVLQNKQKVEFKIDIVLAFQKASIVYKKEGQVFSFPLAFTVEKLTGSG 236
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
166 -----VADYFPEPVTVS----- 177

QY 237 ELWMOABERASSKSMITFDLKNKEVSVKVTVQDRKLGKGLPLHLPLPOLPYOAGSGN 296
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
178 ---WNSGALTSG-----VH-TFPAYL--QSSGLYS 201

QY 297 LTLALBAKTKLHOENVLVVVRATOLQNLTCCEWGFPTSPKMLSLKLENKAVSKREK 356
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
202 LSSVVTYPPSSSLGTQYI-----CNV-----NHKPSNTKVDK 233

QY 357 PVWVLNBPAGMQLLSDSGVLLSNIKVLPTWSTPVPCAPAPKSCDKTHTC-----P 411
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
234 RV-----EPKSCDKTHTCPCPAP 252

QY 412 ELLEGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMYWYDGVENNAKTKPR 471
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
253 ELLEGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMYWYDGVENNAKTKPR 312

QY 472 EEOYNSTYRVVSVTVLHODMNLNGEKYCKSNKALPAPIEKTISKAKGQREPOVYTLR 531
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
313 EEOYNSTYRVVSVTVLHODMNLNGEKYCKSNKALPAPIEKTISKAKGQREPOVYTLR 372

QY 532 PSRDELTKQVSLTCLVKGFPSPDIAVEMESNGQPENNYKTPPVLDSDGSFFLYSKLTV 591
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
373 PSRDELTKQVSLTCLVKGFPSPDIAVEMESNGQPENNYKTPPVLDSDGSFFLYSKLTV 432

QY 592 DKSRMOQGNVFSQGVMEHALNHYTQKSLSPG 625
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
433 DKSRMOQGNVFSQGVMEHALNHYTQKSLSPG 466
```

```
RESULT 93
US-09-313-942-7
; Sequence 7, Application US/09313942
; Patent No. 6472179
; GENERAL INFORMATION:
; APPLICANT: REGENERON PHARMACEUTICALS, INC.
```

```
; TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING
; FILE REFERENCE: REG 203-A
; CURRENT APPLICATION NUMBER: US/09/313,942
; CURRENT FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 09/313,942
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/101,858
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-313-942-7
```

```
Query Match 35.5%; Score 1212.5; DB 4; Length 859;
Best Local Similarity 93.1%; Pred. No. 9.3e-88;
Matches 229; Conservative 3; Mismatches 5; Indels 9; Gaps 2;
```

```
QY 389 TWSTPV---PCPAPPEKSCDKTHTC-----PELIGPSVFLFPKPKDTLMISRTPEVT 439
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
607 TFFTTPKFAQGEIISGEBKSCDKTHTCPCPAPBELIGPSVFLFPKPKDTLMISRTPEVT 666

QY 440 CVVVDVSHEDPEVKENMYVGVFNNAKTKPREQVNSTYRVVSVTVLHODMNLNGEKYK 499
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
667 CVVVDVSHEDPEVKENMYVGVFNNAKTKPREQVNSTYRVVSVTVLHODMNLNGEKYK 726

QY 500 CKVSNKALPAPIEKTISKAKGQREPOVYTLPSRDELTKQVSLTCLVKGFPSPDIAVE 559
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
727 CKVSNKALPAPIEKTISKAKGQREPOVYTLPSRDELTKQVSLTCLVKGFPSPDIAVE 786

QY 560 MESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRMOQGNVFSQGVMEHALNHYTQKS 619
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
787 MESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRMOQGNVFSQGVMEHALNHYTQKS 846

QY 620 LSLSPG 625
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
847 LSLSPG 852
```

```
RESULT 94
US-09-313-942-9
; Sequence 9, Application US/09313942
; Patent No. 6472179
; GENERAL INFORMATION:
; APPLICANT: REGENERON PHARMACEUTICALS, INC.
; TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING
; FILE REFERENCE: REG 203-A
; CURRENT APPLICATION NUMBER: US/09/313,942
; CURRENT FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 09/313,942
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/101,858
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-313-942-9
```

```
Query Match 35.5%; Score 1212.5; DB 4; Length 951;
Best Local Similarity 97.8%; Pred. No. 1.1e-87;
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
```

```
QY 400 EPKSCDKTHTC-----PELIGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 454
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
720 EPKSCDKTHTCPCPAPBELIGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 779
```

```
QY 455 NMVVDGEVHNNAKTKPREEOYNSTYRVSVLTVLHODMLNGKREKCKVSKALPAPIEKT 514
| | | | |
DB 780 NMVVDGEVHNNAKTKPREEOYNSTYRVSVLTVLHODMLNGKREKCKVSKALPAPIEKT 839
QY 515 ISKAKQPREPOVYTLPPSHDELTKNOVSLTCLVKGFPSDIAVEMESNQPENNYKTP 574
| | | | |
DB 840 ISKAKQPREPOVYTLPPSHDELTKNOVSLTCLVKGFPSDIAVEMESNQPENNYKTP 899
QY 575 PVLDSGSPFLYSKLTVDKSRMOQGVFSCSVHGEALHNYTOKSLSLSPG 625
| | | | |
DB 900 PVLDSGSPFLYSKLTVDKSRMOQGVFSCSVHGEALHNYTOKSLSLSPG 950
```

RESULT 95

```
US-09-247-352-3
; Sequence 3, Application US/09247352
; GENERAL INFORMATION:
; APPLICANT: Aruffo, Alejandro A.
; APPLICANT: Siadak, Anthony W.
; APPLICANT: Berry, Karen K.
; APPLICANT: Harris, Linda
; APPLICANT: Thorne, Barbara A.
; APPLICANT: Bajorath, Jurgen
; APPLICANT: Huse, William D.
; APPLICANT: Wu, Herren
; APPLICANT: Watkins, Jeffrey D.
; TITLE OF INVENTION: ANTIBODIES AGAINST HUMAN CD40
; FILE REFERENCE: DB2a SEQUENCE
; CURRENT APPLICATION NUMBER: US/09/247,352
; EARLIER FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 09/026,291
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Human and Mouse
US-09-247-352-3
```

```
Query Match 35.5%; Score 1212; DB 4; Length 451;
Best Local Similarity 46.7%; Pred. No. 4e-88;
Matches 286; Conservative 34; Mismatches 103; Indels 190; Gaps 17;
QY 30 LKKKGDVTLCTASQKSIQFHWNKSNQIKILGNQGSFLTKGSPKLNDRADSRRLMD- 88
| | | | |
DB 11 LKKKGFTVIRISCKAS--GYAFTTGMQWQEMPGKG--LKWIGMINTHSGVPKYVEDF 64
QY 89 QGNFP-----LITIKLIKEDSDTYICEVEDQKEVOULVFGLTANSDTHLLOQ 137
| | | | |
DB 65 KGRFASLETSANTAYLQISNLKNEDTATYFC-VRSNGNAYDLAYFA-----YWGQ 114
QY 138 SLTLTLESPPGSPSPVQCRSPRGKNIQGGKTLVSQLELDSDGTWCTCYLQONKQVEFKI 197
| | | | |
DB 115 GLTVTVAASATKGSVFPLAPSSKSTSG--TALAGCL----- 150
QY 198 DIYVLAFOKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMWQAERASSSKSWITFDLK 257
| | | | |
DB 151 -----VKDYFPEPVTVS-----WNSGALTSG----- 171
QY 258 NKEVSVKRVYODPKLQMGKKLPLHLTLPOALPOYAGSGLTLALEATGKLHQBENVLVVM 317
| | | | |
DB 172 -----VH-TFPAVL-QSSGLYSLSSVTVTPSSSLGTQYTI--- 204
QY 318 RATOLQKNLTCCEVWGPTSPKMLSLKLENKAKVSKREKPVWVLNPEAGMWQCLSDSGQ 377
| | | | |
DB 205 -----CNV-----NHKPSNTKYDKV----- 220
QY 378 VLESNIKVLPTWSTVPCPAPRPKSCDKTHTC-----PELLGGPSVFLPFPKQDTLMI 432
| | | | |
DB 221 -----EPKSCDKTHTCPCPAPRPKSGSVFLPFPKQDTL-I 257
```

```
QY 433 SRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNNAKTKPREEOYNSTYRVSVLTVLHODM 492
| | | | |
DB 258 SRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNNAKTKPREEOYNSTYRVSVLTVLHODM 317
QY 493 LNKKEYCKRVSKNALPAPIEKTISKAKQPREPOVYTLPPSHDELTKNOVSLTCLVKGFY 552
| | | | |
DB 318 LNKKEYCKRVSKNALPAPIEKTISKAKQPREPOVYTLPPSHDELTKNOVSLTCLVKGFY 377
QY 553 PBDIAVEMESNQPENNYKTPPVLDSGSPFLYSKLTVDKSRMOQGVFSCSVHGEALH 612
| | | | |
DB 378 PBDIAVEMESNQPENNYKTPPVLDSGSPFLYSKLTVDKSRMOQGVFSCSVHGEALH 437
QY 613 NHYTOKSLSLSPG 625
| | | | |
DB 438 NHYTOKSLSLSPG 450
```

RESULT 96

```
US-09-466-635-3
; Sequence 3, Application US/09466635
; Patent No. 643516
; GENERAL INFORMATION:
; APPLICANT: Aruffo, Alejandro A.
; APPLICANT: Siadak, Anthony W.
; APPLICANT: Berry, Karen K.
; APPLICANT: Harris, Linda
; APPLICANT: Thorne, Barbara A.
; APPLICANT: Bajorath, Jurgen
; TITLE OF INVENTION: ANTIBODIES AGAINST HUMAN CD40
; FILE REFERENCE: DB2 SEQUENCE
; CURRENT APPLICATION NUMBER: US/09/466,635
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Human and Mouse
US-09-466-635-3
```

```
Query Match 35.5%; Score 1212; DB 4; Length 451;
Best Local Similarity 46.7%; Pred. No. 4e-88;
Matches 286; Conservative 34; Mismatches 103; Indels 190; Gaps 17;
QY 30 LKKKGDVTLCTASQKSIQFHWNKSNQIKILGNQGSFLTKGSPKLNDRADSRRLMD- 88
| | | | |
DB 11 LKKKGFTVIRISCKAS--GYAFTTGMQWQEMPGKG--LKWIGMINTHSGVPKYVEDF 64
QY 89 QGNFP-----LITIKLIKEDSDTYICEVEDQKEVOULVFGLTANSDTHLLOQ 137
| | | | |
DB 65 KGRFASLETSANTAYLQISNLKNEDTATYFC-VRSNGNAYDLAYFA-----YWGQ 114
QY 138 SLTLTLESPPGSPSPVQCRSPRGKNIQGGKTLVSQLELDSDGTWCTCYLQONKQVEFKI 197
| | | | |
DB 115 GLTVTVAASATKGSVFPLAPSSKSTSG--TALAGCL----- 150
QY 198 DIYVLAFOKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMWQAERASSSKSWITFDLK 257
| | | | |
DB 151 -----VKDYFPEPVTVS-----WNSGALTSG----- 171
QY 258 NKEVSVKRVYODPKLQMGKKLPLHLTLPOALPOYAGSGLTLALEATGKLHQBENVLVVM 317
| | | | |
DB 172 -----VH-TFPAVL-QSSGLYSLSSVTVTPSSSLGTQYTI--- 204
QY 318 RATOLQKNLTCCEVWGPTSPKMLSLKLENKAKVSKREKPVWVLNPEAGMWQCLSDSGQ 377
| | | | |
DB 205 -----CNV-----NHKPSNTKYDKV----- 220
QY 378 VLESNIKVLPTWSTVPCPAPRPKSCDKTHTC-----PELLGGPSVFLPFPKQDTLMI 432
| | | | |
DB 221 -----EPKSCDKTHTCPCPAPRPKSGSVFLPFPKQDTL-I 257
QY 433 SRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNNAKTKPREEOYNSTYRVSVLTVLHODM 492
```

Db 258 SRTEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKRREQDYNSTYRVSVLTVLDHQM 317
Qy 493 LINGEYKCKVKNKALPAPIETKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGF 552
Db 318 LINGEYKCKVKNKALPAPIETKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGF 377
Qy 553 PSDIAVESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHEALH 612
Db 378 PSDIAVESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHEALH 437
Qy 613 NHYTQKSLISLSPG 625
Db 438 NHYTQKSLISLSPG 450
RESULT 97
US-09-301-593-30
Sequence 30, Application US/09301593A
Patent No. 6455677
GENERAL INFORMATION:
APPLICANT: Park, John E.
APPLICANT: Garin-Chesa, Pilar
APPLICANT: Bamberger, Uwe
APPLICANT: Legier, Olivier
APPLICANT: Saidana, Jose M.
APPLICANT: Rettig, Wolfgang J.
TITLE OF INVENTION: RAP-specific Antibody with Improved Producibility
FILE REFERENCE: 0652.1890001
CURRENT APPLICATION NUMBER: US/09/301.593A
CURRENT FILING DATE: 1999-04-29
EARLIER APPLICATION NUMBER: EP 98107925.4
EARLIER FILING DATE: 1998-04-30
EARLIER APPLICATION NUMBER: US 60/086.049
EARLIER FILING DATE: 1998-05-18
NUMBER OF SEQ ID NOS: 108
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 30
LENGTH: 472
TYPE: PROT
ORGANISM: Homo sapiens
US-09-301-593-30
Query Match 35.5%; Score 1212; DB 4; Length 472;
Best Local Similarity 49.4%; Pred. No. 4.3e-88;
Matches 266; Conservative 39; Mismatches 74; Indels 160; Gaps 13;
Qy 147 PGSSPSVQCRSPR-----GKNIOG-----G 166
Db 33 PGASVKNKCKTSRTFTFTYTIHWVRQSHKSLWIGINPNNGIPIYNQKFKGRATLTVG 92
Qy 167 KTLVSQLEL-----QDSGTWTCVLOQKVEFKIDIVLAFOKASSIYKKEGBOYEF 221
Db 93 KSSSTAYVELRSLTSEDSAVYFCA---RRRIAYGYD-----EGHMDY 132
Qy 222 SFPLAFYEXKLTGSELWMOABRASSSKSWITPDLKNEVSVKRVTQDPKLOMGKULPLH 281
Db 133 -----WGQTSVTSSS-----TKGSPVFPPLAPSSKTSQGTALG 168
Qy 282 LTLPOALPOYA-----GSGNLTILALEA-----KTGKLHQBVLVVMRATQL-QKNLTCEVW 331
Db 169 CLVVDYFPEPVTAVNSGALTSVCHTTPPAVLQSSGLYSVTVYVSSLSGTQYIYICV- 227
Qy 332 GPSPKMLSLKLENKAKVSKREKPVAVLPEAGMOCCLISDGOVLLESNIKVLPTWS 391
Db 228 -----NHKPSNTKVDKCV----- 240
Qy 392 TPVCPAPPEKSCDKTHC-----PELLGSPVFLPPKPKDTLMISTPEVTCVVVDVS 446
Db 241 -----EPKSCDKTHCPCPPAPPELLGSPSVFLPPKPKDTLMISTPEVTCVVVDVS 292
Qy 447 HEDPEVKFNWYVDGVEVHNAKTKRREQDYNSTYRVSVLTVLDHQM LINGEYKCKVKNK 506
|||||

Db 293 HEDPEVKFNWYVDGVEVHNAKTKRREQDYNSTYRVSVLTVLDHQM LINGEYKCKVKNK 352
Qy 507 LPAPIETKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVESNGQPE 566
Db 353 LPAPIETKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVESNGQPE 412
Qy 567 ENNYKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHEALHNYHTQKSLISLSPG 625
Db 413 ENNYKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHEALHNYHTQKSLISLSPG 471
RESULT 98
US-08-595-043A-50
Sequence 50, Application US/08595043A
Patent No. 5935824
GENERAL INFORMATION:
APPLICANT: SGARLATO, GREGORY D.
TITLE OF INVENTION: PROTEIN EXPRESSION SYSTEM
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/595.043A
FILING DATE: 31-JAN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: SGAR-00371
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 232 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-595-043A-50
Query Match 35.5%; Score 1211.5; DB 2; Length 232;
Best Local Similarity 97.8%; Pred. No. 1.7e-88;
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
Qy 400 EPKSCDKTHC-----PELLGSPVFLPPKPKDTLMISTPEVTCVVVDVSHEDPEVKF 454
Db 1 EPKSCDKTHCPCPPAPPELLGSPSVFLPPKPKDTLMISTPEVTCVVVDVSHEDPEVKF 60
Qy 455 NMVYDGEVHNAKTKRREQDYNSTYRVSVLTVLDHQM LINGEYKCKVKNKALPAPIETKT 514
Db 61 NMVYDGEVHNAKTKRREQDYNSTYRVSVLTVLDHQM LINGEYKCKVKNKALPAPIETKT 120
Qy 515 ISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVESNGQPENNYKTPP 574
Db 121 ISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVESNGQPENNYKTPP 180
Qy 575 PVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHEALHNYHTQKSLISLSPG 625
Db 181 PVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVMHEALHNYHTQKSLISLSPG 231
RESULT 99
US-09-178-869-2

```

; Sequence 2, Application US/09178869B
; Patent No. 6197294
; GENERAL INFORMATION:
; APPLICANT: Tao, Weng
; APPLICANT: Wong, Shou
; APPLICANT: Hickey, William F.
; APPLICANT: Hamman, Joseph P.
; APPLICANT: Baetge, E. Edward
; TITLE OF INVENTION: CELL SURFACE-INDUCED MACROPHAGE ACTIVATION
; FILE REFERENCE: 17810-043
; CURRENT APPLICATION NUMBER: US/09/178,869B
; CURRENT FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-178-869-2

```

```

Query Match      35.5%; Score 1211.5; DB 3; Length 331;
Best Local Similarity 97.8%; Pred. No. 2.8e-88;
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

```

```

QY 400 EPKSCDKTHTC-----PELLGSPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 454
DB 100 EPKSCDKTHTCPAPABELLGSPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 159
QY 455 NMVYDGEVYHNAKTKPREEOYNSTYRVSVLTVLHODMLNGKEYCKCKVSNKALPAPIEKT 514
DB 160 NMVYDGEVYHNAKTKPREEOYNSTYRVSVLTVLHODMLNGKEYCKCKVSNKALPAPIEKT 219
QY 515 ISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTT 574
DB 220 ISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTT 279
QY 575 PVLDSGSEFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 625
DB 280 PVLDSGSEFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 330

```

```

RESULT 100
US-09-761-413-2
; Sequence 2, Application US/09761413
; Patent No. 6506891
; GENERAL INFORMATION:
; APPLICANT: Tao, Weng
; APPLICANT: Wong, Shou
; APPLICANT: Hickey, William F.
; APPLICANT: Hamman, Joseph P.
; APPLICANT: Baetge, E. Edward
; TITLE OF INVENTION: CELL SURFACE-INDUCED MACROPHAGE ACTIVATION
; FILE REFERENCE: 17810-043
; CURRENT APPLICATION NUMBER: US/09/761,413
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US/09/178,869
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-761-413-2

```

```

Query Match      35.5%; Score 1211.5; DB 4; Length 331;
Best Local Similarity 97.8%; Pred. No. 2.8e-88;
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

```

```

QY 400 EPKSCDKTHTC-----PELLGSPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 454
DB 100 EPKSCDKTHTCPAPABELLGSPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 159

```

```

QY 455 NMVYDGEVYHNAKTKPREEOYNSTYRVSVLTVLHODMLNGKEYCKCKVSNKALPAPIEKT 514
DB 160 NMVYDGEVYHNAKTKPREEOYNSTYRVSVLTVLHODMLNGKEYCKCKVSNKALPAPIEKT 219
QY 515 ISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTT 574
DB 220 ISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTT 279
QY 575 PVLDSGSEFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 625
DB 280 PVLDSGSEFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 330

```

```

Search completed: August 3, 2004, 13:17:11
Job time : 29.0695 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:50:09 ; Search time 65.926 Seconds
(without alignments)
2777.216 Million cell updates/sec

Title: SEQ7
Perfect score: 3414
Sequence: 1 MNRGVPRHLLVLQLALP.....DETCAEQDELGLWTTDP 648

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 125 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3229.5	94.6	631 1 AAP93009	AAP93009 Genetic C
2	3229.5	94.6	631 3 AAB19508	AAB19508 CD4-IgG1
3	3229.5	94.6	631 3 AAY51079	AAY51079 Human FUS
4	3229.5	94.6	631 3 AAY59169	AAY59169 CD4-Ig fu
5	3176.5	93.0	729 1 AAP93008	AAP93008 Genetic C
6	3176.5	93.0	729 3 AAB19507	AAB19507 CD4-IgG1
7	3176.5	93.0	729 3 AAY59168	AAY59168 CD4-Ig fu
8	3166.5	92.8	729 3 AAY51078	AAY51078 Human FUS
9	3163	92.6	616 3 AAY51082	AAY51082 Human FUS
10	3163	92.6	616 3 AAY59172	AAY59172 CD4-Ig fu
11	3132	91.7	614 1 AAP93012	AAP93012 Genetic C
12	3071	90.0	616 3 AAB19511	AAB19511 CD4-IgG1
13	2212	64.8	534 2 AAR26531	AAR26531 Sequence
14	2140	62.7	435 2 AAR26530	AAR26530 Sequence
15	2116	62.0	530 2 AAR26783	AAR26783 CD4-IgG2
16	2116	62.0	530 3 AAY85080	AAY85080 CD4-IgG2
17	2116	62.0	530 4 AAB67323	AAB67323 CD4-IgG2
18	2116	62.0	530 4 AAB80884	AAB80884 Human CD4
19	2116	62.0	530 6 ABG71123	ABG71123 CD4-Igmun
20	2107.5	61.7	450 6 AAE37574	AAE37574 Human DID
21	2106	61.7	530 6 AAR46679	AAR46679 CD4-IgG2
22	2088	61.2	449 6 AAE37576	AAE37576 Human FDI
23	2073	60.7	481 3 AAB19510	AAB19510 CD4-IgM f
24	2073	60.7	481 3 AAY59171	AAY59171 CD4-Ig fu
25	2067	60.5	481 3 AAY51081	AAY51081 Human FUS

26	2065	60.5	481 1 AAP93011	AAP93011 Genetic C
27	2059	60.3	432 2 AAR26782	AAR26782 CD4-gamma
28	2059	60.3	432 2 AAR46678	AAR46678 CD4-gamma
29	2059	60.3	432 2 AAY85079	AAY85079 Human CD4
30	2059	60.3	432 4 AAB67322	AAB67322 CD4-gamma
31	2059	60.3	432 4 AAB80883	AAB80883 Human CD4
32	2059	60.3	432 6 ABG71122	ABG71122 CD4-gamma
33	2046.5	59.9	436 3 AAY51080	AAY51080 Human FUS
34	2041	59.8	474 3 AAY59170	AAY59170 CD4-Ig fu
35	2036.5	59.7	532 2 AAR27278	AAR27278 CD4-gamma
36	2036.5	59.7	532 2 AAR78678	AAR78678 T-cell re
37	2036.5	59.7	532 2 AAR89458	AAR89458 CD4-eta f
38	2036.5	59.7	532 2 AAB83142	AAB83142 Chimeric
39	2036.5	59.7	575 2 AAR27276	AAR27276 CD4-zeta
40	2036.5	59.7	575 2 AAR78676	AAR78676 T-cell re
41	2036.5	59.7	575 2 AAR89456	AAR89456 CD4-zeta
42	2036.5	59.7	575 2 AAW02213	AAW02213 CD4-T-cell
43	2036.5	59.5	575 2 AAW83140	AAW83140 Chimeric
44	2032.5	59.5	462 2 AAR27277	AAR27277 CD4-eta p
45	2032.5	59.5	462 2 AAR78677	AAR78677 T-cell re
46	2032.5	59.5	462 2 AAR89457	AAR89457 CD4-gamma
47	2032.5	59.5	462 2 AAW02214	AAW02214 CD4-FC re
48	2032.5	59.5	462 2 AAW83142	AAW83142 Chimeric
49	2032.5	59.5	532 2 AAW02215	AAW02215 CD4-T-cell
50	2029	59.4	398 2 AAR89450	AAR89450 CD4-DL-D4
51	2029	59.4	458 2 AAR07769	AAR07769 DNA encod
52	2026	59.3	398 2 AAR78673	AAR78673 CD4 domai
53	2026	59.3	416 3 AAB19509	AAB19509 CD4-IgM f
54	2021	59.2	400 2 AAR06374	AAR06374 Truncated
55	2021	59.2	458 1 AAB81990	AAB81990 Clone PT4
56	2021	59.2	458 1 AAP91369	AAP91369 T4 protei
57	2021	59.2	458 2 AAY19825	AAY19825 Soluble h
58	2021	59.2	2037 2 AAR04032	AAR04032 Full leng
59	2021	59.2	2050 2 AAR07641	AAR07641 Deduced s
60	2018	59.1	394 1 AAP93506	AAP93506 Derived s
61	2017	59.1	402 1 AAP91922	AAP91922 Sequence
62	2017	59.1	402 1 AAP94757	AAP94757 Sequence
63	2015	59.0	394 3 AAY88328	AAY88328 T4 glycop
64	2015	59.0	458 3 AAY88329	AAY88329 T4 glycop
65	2015	59.0	458 4 AAB81502	AAB81502 Human CD4
66	2015	59.0	458 7 ADD25609	ADD25609 Binding d
67	2015	59.0	458 7 ADE57489	ADE57489 Human Pro
68	2015	59.0	473 7 ADA44807	ADA44807 CD4/TCR C
69	2015	59.0	519 2 AAR20152	AAR20152 Human CD4
70	2013	59.0	458 2 AAR13491	AAR13491 Human CD4
71	2012	58.9	394 2 AAY39825	AAY39825 Soluble h
72	2007	58.8	458 4 AAG79087	AAG79087 Amino aci
73	2001	58.5	394 3 AAB07768	AAB07768 The solub
74	1997	58.5	458 7 ADE65841	ADE65841 Human CD4
75	1993	58.4	458 2 AAR06373	AAR06373 T4 encode
76	1993	58.4	524 1 AAP94703	AAP94703 Sequence
77	1993	58.4	2458 2 AAR07640	AAR07640 Deduced p
78	1993	58.3	2458 2 AAR04031	AAR04031 Full leng
79	1991	58.3	399 2 AAR20151	AAR20151 Chimpanze
80	1991	58.3	458 2 AAR11285	AAR11285 gp120 bin
81	1991	58.3	458 2 AAR10988	AAR10988 Chimpanze
82	1986	58.2	400 2 AAR20150	AAR20150 Chimpanze
83	1985	58.1	458 2 AAR04910	AAR04910 T4 protei
84	1982.5	58.1	399 1 AAP93010	AAP93010 Genetic C
85	1946	57.0	400 2 AAR20154	AAR20154 Sol. rhes
86	1929.5	56.5	729 2 AAR41042	AAR41042 CD4-GBPH
87	1921	56.3	458 2 AAR15149	AAR15149 CD4 coo
88	1908.5	55.9	942 2 AAR41041	AAR41041 CD4-GBP13
89	1908	55.9	384 1 AAP90833	AAP90833 Amino aci
90	1908	55.9	1786 2 AAR41043	AAR41043 CD4-EBA17
91	1904	55.8	434 1 AAP96151	AAP96151 Sequence
92	1904	55.8	434 1 AAP93557	AAP93557 Fusion of
93	1894	55.5	375 2 AAR07721	AAR07721 Recombina
94	1892	55.4	433 2 AAW41376	AAW41376 Human CD4
95	1891	55.4	369 2 AAY39824	AAY39824 Soluble h
96	1891	55.4	369 3 AAY88327	AAY88327 T4 glycop
97	1891	55.4	370 1 AAP93528	AAP93528 Human sol
98	1891	55.4	370 4 AAB83356	AAB83356 Human CD4

```

99 1887 55.3 433 3 AA54500 Aay54500 Amino aci
100 1884 55.2 432 2 AAR74222 Aar74222 Eptlope o
101 1884 55.2 432 2 AAY30514 Aay30514 Predicted
102 1884 55.2 432 6 ADA25188 Ada25188 CD4 epitc
103 1879 55.0 370 2 AAR12956 Aar12956 Lys (64) C
104 1878 55.0 370 2 AAR12954 Aar12954 His (40) C
105 1878 55.0 370 2 AAR12964 Aar12964 Val (143)
106 1878 55.0 370 2 AAR12965 Aar12965 Arg (150)
107 1877 55.0 370 2 AAR12958 Aar12958 Gln (80) C
108 1874 54.9 370 2 AAR12951 Aar12951 Glu (17)
109 1872 54.8 370 2 AAR12957 Aar12957 Asn (72)
110 1871 54.8 370 2 AAR12959 Aar12959 Asn (88)
111 1871 54.7 370 2 AAR12962 Aar12962 Leu (127)
112 1866 54.7 370 2 AAR12966 Aar12966 Thr (162)
113 1865 54.6 370 2 AAR12952 Aar12952 Ile (23)
114 1864 54.6 370 2 AAR12967 Aar12967 Asp (155)
115 1862 54.5 370 2 AAR12955 Aar12955 Gly (48)
116 1861 54.5 370 2 AAR12963 Aar12963 His (132)
117 1861 54.5 370 2 AAR12960 Aar12960 Lys (99)
118 1861 54.5 435 1 AAR90992 Aar90992 Human CD4
119 1858 54.4 370 2 AAR12961 Aar12961 Ser (121)
120 1854 54.3 435 1 AAR12961 Aar12961 Ser (121)
121 1852 54.2 370 2 AAR12959 Aar12959 Sequence
122 1850 54.2 433 2 AAR08335 Aar08335 Thr (27)
123 1849 54.2 400 2 AAR08335 Aar08335 CD4 . 3/20
124 1848 54.1 458 2 AAR10987 Aar10987 Rhesus mo
125 1847 54.1 458 2 AAR20148 Aar20148 Cynomolgus

```

ALIGNMENTS

```

RESULT 1
ID AAP93009 standard; protein; 631 AA.
XX AAP93009;
AC AAP93009;
DT 25-MAR-2003 (revised)
DT 02-NOV-1992 (first entry)
XX
DE Genetic construct which encodes CD4 linked to human IgG1 at the Bsp site
DE upstream of the hinge region (fusion protein CD4E-gamma-1).
XX
XX Fusion protein; immunoglobulin-like molecule; HIV; SIV; therapy;
XX diagnosis; CD4; gp120; binding fragment; glycoprotein; variable region.
XX
XX Homo sapiens.
XX
XX EP325262-A.
XX
XX 26-JUL-1989.
XX
XX 20-JAN-1989; 89EP-00100913.
XX
XX 22-JAN-1988; 88US-00147351.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Seed B;
XX
XX WPI; 1989-214472/30.
XX N-PSDB; AAN90357.
XX
XX Immunoglobulin-CD4 fusion proteins - used for treating HIV or SIV
XX infections or detecting HIV or SIV in sample.
XX
XX Example; Table 2, Page 24-33; 68pp; English.
XX
XX The fusion protein genes of the invention pref. comprise cDNA sequences
XX which encode CD4 or a fragment which binds gp120 ligated to an expression
XX CC plasmid which encodes an antibody in which the variable region of the
XX gene has been deleted (see WO87-02671). The CD4 portion of the fusion

```

protein may comprise the complete CD4 sequence, the 370 AA extracellular region and the membrane spanning domain, or the extracellular region. The Ig heavy chain is pref. from IgM, IgG1 or IgG3. The following are specifically claimed: fusion proteins CD4H-gamma-1, CD4Hmu, CD4Hpm, CD4E-gamma1, and CD4Hmu (No. 67608), pCD4P-gamma-1, and pCD4E-gamma-1 (No. 67610). The plasmid containing (pCD4E-gamma-1) has been deposited in E. coli (MC1061/p3) at the ATCC under accession number 67610. (Updated on 25-MAR-2003 to correct PA field.)

Sequence 631 AA;

Query Match 94.6%; Score 3229.5; DB 1; Length 631;
Best Local Similarity 98.3%; Pred. No. 36-166;
Matches 619; Conservative 0; Mismatches 6; Indels 5; Gaps 1;

```

QY 1 MNRGVFRLHLVLTALPAATGKVVLAGKGDVELTCTASQKSIQFMKNSQIK 60
DB 1 MNRGVFRLHLVLTALPAATGKVVLAGKGDVELTCTASQKSIQFMKNSQIK 60
QY 61 IIGNGSFLTKGSPKLNDRADSRRLMDGNFPLIKNLIKIDSDTYICEVEDQKEEVL 120
DB 61 IIGNGSFLTKGSPKLNDRADSRRLMDGNFPLIKNLIKIDSDTYICEVEDQKEEVL 120
QY 121 LVFGLTANSDDLQGSLLTLTLESPPGSSPEVQCSPPGKNIQGGKTLVSQLELDSDG 180
DB 121 LVFGLTANSDDLQGSLLTLTLESPPGSSPEVQCSPPGKNIQGGKTLVSQLELDSDG 180
QY 181 TWCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGOVEFSPLAFTYEKLTGSGELMW 240
DB 181 TWCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGOVEFSPLAFTYEKLTGSGELMW 240
QY 241 QAEKSSSKSWITFDLKNKEVSKVKTQDPKLOMGKPLHLTLPOLQVYSGSLTLTA 300
DB 241 QAEKSSSKSWITFDLKNKEVSKVKTQDPKLOMGKPLHLTLPOLQVYSGSLTLTA 300
QY 301 LEAKTGKLEQEVNLYVMRATOLQKMLTCEWGPSPKMLSLKLENKAKVSKREKPVW 360
DB 301 LEAKTGKLEQEVNLYVMRATOLQKMLTCEWGPSPKMLSLKLENKAKVSKREKPVW 360
QY 361 LNPEAGMOCCLSDSGQVLLBSNIXLPTWSTPVPVCPADPEPKSCDKTHTC-----DELLG 415
DB 361 LNPEAGMOCCLSDSGQVLLBSNIXLPTWSTPVPVCPADPEPKSCDKTHTC-----DELLG 415
QY 416 GSVFLPFPKPKDPTLMISRPTEVTCVVDVSHEDPEVKVMYVDGVEVNAKTKPREEO 475
DB 416 GSVFLPFPKPKDPTLMISRPTEVTCVVDVSHEDPEVKVMYVDGVEVNAKTKPREEO 475
QY 476 NSTYRVSVLTVLHODMLNGKEYKCVSNKALPAPIEKTISKAKGPREPOVYTLPPSRD 535
DB 476 NSTYRVSVLTVLHODMLNGKEYKCVSNKALPAPIEKTISKAKGPREPOVYTLPPSRD 535
QY 481 NSTYRVSVLTVLHODMLNGKEYKCVSNKALPAPIEKTISKAKGPREPOVYTLPPSRD 540
DB 481 NSTYRVSVLTVLHODMLNGKEYKCVSNKALPAPIEKTISKAKGPREPOVYTLPPSRD 540
QY 536 ELTKNOVSLTCLVKGFPSPDIAVEWESNQPENNYTTTPVDLSDGSFLYSKLTVDKSR 595
DB 536 ELTKNOVSLTCLVKGFPSPDIAVEWESNQPENNYTTTPVDLSDGSFLYSKLTVDKSR 595
QY 541 ELTKNOVSLTCLVKGFPSPDIAVEWESNQPENNYTTTPVDLSDGSFLYSKLTVDKSR 600
DB 541 ELTKNOVSLTCLVKGFPSPDIAVEWESNQPENNYTTTPVDLSDGSFLYSKLTVDKSR 600
QY 596 MOGNVFSQVWHEALHNHYTQKSLSPG 625
DB 601 MOGNVFSQVWHEALHNHYTQKSLSPG 630

```

RESULT 2

AAB19508 standard; protein; 631 AA.

AC AAB19508;

DT 09-JAN-2001 (first entry)

XX CD4-IgG1 fusion protein CH4Egamma1.

XX CD4; IgG1; human; CD4Egamma1; fusion protein; immunoglobulin; HIV; SIV;
XX gp120; therapy; diagnosis.

```

OS Homo sapiens.
XX Key Location/Qualifiers
XX Protein 1..395
FT /note="CD4 extracellular region"
FT Protein 400..631
FT /note="IgG1 heavy chain"
XX US6117656-A.
XX 12-SEP-2000.
XX 07-JUN-1995; 95US-00479353.
XX 22-JAN-1988; 88US-00147351.
XX 23-JAN-1989; 89US-00289586.
XX 09-JUN-1992; 92US-00896781.
XX 12-APR-1993; 93US-00057952.
XX 04-FEB-1994; 94US-00191708.
XX (GENO ) GEN HOSPITAL CORP.
XX Seed B;
XX WPI; 2000-586558/55.
XX N-PSDB; AAA50661.
XX CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or
XX SIV.
XX Example 1; Col 29-42; 39pp; English.
XX The present sequence is that of fusion protein CD4Egammal comprising the
XX extracellular portion of CD4, which binds to HIV gp120, linked at its C-
XX terminus to the human IgG1 heavy chain. To obtain the fusion protein, DNA
XX encoding CD4 was linked to IgG1 DNA at the Esp site upstream of the hinge
XX region (see AAA50661). Fusion protein CD4Egammal and a nucleic acid
XX encoding it are claimed. Also claimed are a vector comprising the nucleic
XX acid, and a method of producing the fusion protein in secreted form using
XX a transformed host cell. The fusion protein may further comprise a
XX therapeutic agent, radiolabel or NMR imaging agent. The fusion protein
XX can be administered to an animal (including humans) for treatment of HIV
XX or SIV infection, and can also be used in assays for HIV or SIV, imaging
XX and tissue stains. IgG1 fusion proteins such as CD4Egammal provide both
XX complement-mediated and cell-mediated immunity.
XX Sequence 631 AA:
XX
XX Query Match 94.6%; Score 3229.5; DB 3; Length 631;
XX Best Local Similarity 98.3%; Pred. No. 3e-166;
XX Matches 619; Conservative 0; Mismatches 6; Indels 5; Gaps 1;
XX
XX 1 MNRGVPFRHLLLVQLALLPAATQGNKVVLGKGGTVLCTCTASQKSIQPHMKNNOIK 60
XX 1 MNRGVPFRHLLLVQLALLPAATQGNKVVLGKGGTVLCTCTASQKSIQPHMKNNOIK 60
XX 1 ILGNQGSFLTKGPKSLNDRADSRSLMDOGFPFLIKKLTIEDSTYICEVEDQKEVQL 120
XX 61 ILGNQGSFLTKGPKSLNDRADSRSLMDOGFPFLIKKLTIEDSTYICEVEDQKEVQL 120
XX 61 ILGNQGSFLTKGPKSLNDRADSRSLMDOGFPFLIKKLTIEDSTYICEVEDQKEVQL 120
XX 121 LVFGITANSPTHLQOGSLTLTLESPGSSPVOCRSRGNIOGKTLVSQLELDOSG 180
XX 121 LVFGITANSPTHLQOGSLTLTLESPGSSPVOCRSRGNIOGKTLVSQLELDOSG 180
XX 121 LVFGITANSPTHLQOGSLTLTLESPGSSPVOCRSRGNIOGKTLVSQLELDOSG 180
XX 181 TWTCITVLONQKVEFKIDIVLAPQKASSIYKKEGEVFSFPLAFVETKLTSGGELMW 240
XX 181 TWTCITVLONQKVEFKIDIVLAPQKASSIYKKEGEVFSFPLAFVETKLTSGGELMW 240
XX 181 TWTCITVLONQKVEFKIDIVLAPQKASSIYKKEGEVFSFPLAFVETKLTSGGELMW 240
XX 241 QAERASSSSKSWITPDLKKEVSVKRVTPDPKLGKGLPLHLTLTPOALPOYAGSGNTLTA 300
XX 241 QAERASSSSKSWITPDLKKEVSVKRVTPDPKLGKGLPLHLTLTPOALPOYAGSGNTLTA 300
XX 301 LEATGKLGHEVNLVVMRATQLOKNLTCEVWGPTSPKLMJSLKLENKAQVSKREKPVWV 360

```

```

DB 301 LEATGKLGHEVNLVVMRATQLOKNLTCEVWGPTSPKLMJSLKLENKAQVSKREKPVWV 360
XX 361 LNPEAGMWOCLLSSGOVLLSNTKVLPTWSTPVPCPAPEKSCDKTHTC-----PEILG 415
XX 361 LNPEAGMWOCLLSSGOVLLSNTKVLPTWSTPVHADPEEPKSCDKTHTCPPCPAPELLG 420
XX 416 GPSVFLPPPKDMLTSTRPEVTCVAVDVSHEDPEVKNYVGVVHNAKTPREEROY 475
XX 421 GPSVFLPPPKDMLTSTRPEVTCVAVDVSHEDPEVKNYVGVVHNAKTPREEROY 480
XX 476 NSTRVVSVLTVLHQDMLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRD 535
XX 481 NSTRVVSVLTVLHQDMLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRD 540
XX 536 ELTKNQVSLTCLVKGFIPPSDIAVEESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSR 595
XX 541 ELTKNQVSLTCLVKGFIPPSDIAVEESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSR 600
XX 596 WQGNVFSQSVMEALHNHYTQKSLSPG 625
XX 601 WQGNVFSQSVMEALHNHYTQKSLSPG 630
XX
XX RESULT 3
XX AAY51079
XX ID AAY51079 standard; protein; 631 AA.
XX AC AAY51079;
XX DT 23-MAR-2000 (first entry)
XX DE Human fusion protein CD4Egammal.
XX KW Fusion protein; human; CD4; IgG1; immunoglobulin; gp120;
XX KW anti-human immunodeficiency virus; CD4Egammal.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN US6004781-A.
XX PD 21-DEC-1999.
XX PP 04-FEB-1994; 94US-00191708.
XX PR 22-JAN-1988; 88US-00147351.
XX PR 23-JAN-1989; 89US-00289586.
XX PR 09-JUN-1992; 92US-00896781.
XX PR 12-APR-1993; 93US-00057952.
XX PA (GENO ) GEN HOSPITAL CORP.
XX PI Seed B;
XX WPI; 2000-085792/07.
XX N-PSDB; AAZ44062.
XX Fusion protein useful for the treatment of human immunodeficiency virus.
XX Example 1; Col 29-42; 39pp; English.
XX This invention describes a novel nucleic acid (I) encoding a fusion
XX protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
XX and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
XX chain (III). The products of the invention have anti-human
XX immunodeficiency virus (HIV) activity and are capable of binding to
XX gp120. The fusion protein is useful for treating human immunodeficiency
XX virus (HIV) or simian immunodeficiency virus (SIV). This sequence
XX represents the fusion protein CD4Egammal which is constructed from CD4
XX linked to human IgG1 upstream of the hinge region
XX Sequence 631 AA:

```

Query Match 94.6%; Score 3229.5; DB 3; Length 631;
 Best Local Similarity 98.3%; Pred. No. 3e-166;
 Matches 619; Conservative 0; Mismatches 6; Indels 5; Gaps 1;

QY 1 MNRGVPRHLLVLTQALPAAATQGNKVVLGKKGDTVELTCTASQKKSIOFHKNSNOIK 60
 |||||
 DB 1 MNRGVPRHLLVLTQALPAAATQGNKVVLGKKGDTVELTCTASQKKSIOFHKNSNOIK 60

QY 61 ILGNQGSFLTKGSPKLNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVOL 120
 |||||
 DB 61 ILGNQGSFLTKGSPKLNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVOL 120

QY 121 LVFGLTANSDTHLLOQGSJLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 |||||
 DB 121 LVFGLTANSDTHLLOQGSJLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 181 TWICTVLQONQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240
 |||||
 DB 181 TWICTVLQONQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240

QY 241 QAERASSKSWITFDLKNKEVSXKRVYQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
 |||||
 DB 241 QAERASSKSWITFDLKNKEVSXKRVYQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300

QY 301 LEAKTGKHOEVNVLVVRATQLOKNTCEVWGPTSPKMLSLKENKAVSKREKRVWY 360
 |||||
 DB 301 LEAKTGKHOEVNVLVVRATQLOKNTCEVWGPTSPKMLSLKENKAVSKREKRVWY 360

QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEAPPEPKSCDKTHTC-----PELLG 415
 |||||
 DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEAPPEPKSCDKTHTC-----PELLG 415

QY 416 GPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 475
 |||||
 DB 416 GPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 475

QY 476 NSTYRVSVLTVLHODPLNGKEYKCKVSNKALPAPIEKTISKAKGOREQVYVTLPPSRD 535
 |||||
 DB 476 NSTYRVSVLTVLHODPLNGKEYKCKVSNKALPAPIEKTISKAKGOREQVYVTLPPSRD 535

QY 536 ELTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSR 595
 |||||
 DB 536 ELTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSR 595

QY 596 WQGNVFSQSVMEHALHNHYTQKSLSPG 625
 |||||
 DB 596 WQGNVFSQSVMEHALHNHYTQKSLSPG 630

RESULT 4
 ID AAY59169 standard; protein; 631 AA.
 XX
 AC AAY59169;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE CD4-Ig fusion protein CD4Egammal.
 XX
 KM HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;
 KM secreted protein; SIV infection; medicament.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN CAL340741-C.
 XX
 PD 14-SEP-1999.
 XX
 PF 20-JAN-1989; 89CA-00588749.
 XX
 PR 20-JAN-1989; 89CA-00588749.

XX (GENE) GEN HOSPITAL CORP.
 PA Seed B;
 XX
 XX WPI; 2000-063015/06.
 DR N-PSDB; A4248202.
 PT New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in
 XX the treatment of HIV or simian immunodeficiency virus infections.
 XX
 XX Example 1; Page 37-46; 89pp; English.

CC The invention provides a fusion gene encoding a fusion protein that
 CC comprises an extracellular CD4 DNA sequence or its fragment which binds
 CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA
 CC sequence of an Ig heavy or light chain, where the DNA sequence encoding
 CC the variable region has been replaced with the DNA sequence which encodes
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is
 CC capable of being secreted. The fusion proteins are useful for treating
 CC HIV or SIV infections in animals, preferably humans. They are also useful
 CC for producing medicaments which can be used for treating HIV or SIV
 CC infections in humans. The present sequence represents the fusion protein
 CC CD4Egammal where the CD4 is linked to human IgG1 at the Bsp site upstream
 CC of the hinge region

XX Sequence 631 AA;
 SQ

Query Match 94.6%; Score 3229.5; DB 3; Length 631;
 Best Local Similarity 98.3%; Pred. No. 3e-166;
 Matches 619; Conservative 0; Mismatches 6; Indels 5; Gaps 1;

QY 1 MNRGVPRHLLVLTQALPAAATQGNKVVLGKKGDTVELTCTASQKKSIOFHKNSNOIK 60
 |||||
 DB 1 MNRGVPRHLLVLTQALPAAATQGNKVVLGKKGDTVELTCTASQKKSIOFHKNSNOIK 60

QY 61 ILGNQGSFLTKGSPKLNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVOL 120
 |||||
 DB 61 ILGNQGSFLTKGSPKLNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVOL 120

QY 121 LVFGLTANSDTHLLOQGSJLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
 |||||
 DB 121 LVFGLTANSDTHLLOQGSJLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 181 TWICTVLQONQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240
 |||||
 DB 181 TWICTVLQONQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240

QY 241 QAERASSKSWITFDLKNKEVSXKRVYQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
 |||||
 DB 241 QAERASSKSWITFDLKNKEVSXKRVYQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300

QY 301 LEAKTGKHOEVNVLVVRATQLOKNTCEVWGPTSPKMLSLKENKAVSKREKRVWY 360
 |||||
 DB 301 LEAKTGKHOEVNVLVVRATQLOKNTCEVWGPTSPKMLSLKENKAVSKREKRVWY 360

QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEAPPEPKSCDKTHTC-----PELLG 415
 |||||
 DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEAPPEPKSCDKTHTC-----PELLG 415

QY 416 GPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 475
 |||||
 DB 416 GPSVFLFPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 475

QY 476 NSTYRVSVLTVLHODPLNGKEYKCKVSNKALPAPIEKTISKAKGOREQVYVTLPPSRD 535
 |||||
 DB 476 NSTYRVSVLTVLHODPLNGKEYKCKVSNKALPAPIEKTISKAKGOREQVYVTLPPSRD 535

QY 536 ELTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSR 595
 |||||
 DB 536 ELTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSR 595

QY 596 WQGNVFSQSVMEHALHNHYTQKSLSPG 625
 |||||
 DB 596 WQGNVFSQSVMEHALHNHYTQKSLSPG 625

```

Db      601  WQGNVVFSCSVHNEALHNHYTQKSLSPG 630
|||||
RESULT 5
AAP93008
ID      AAP93008 standard; protein; 729 AA.
XX
XX      AAP93008;
AC
XX      25-MAR-2003 (revised)
DT      02-NOV-1992 (first entry)
XX
DE      Genetic construct which encodes CD4 linked to human IgG1 at the Hind3
XX      site upstream of the CH1 region (fusion protein CD4H-gamma-1).
XX      Fusion protein; immunoglobulin-like molecule; HIV; SIV; therapy;
XX      diagnosis; CD4; gp120; binding fragment; glycoprotein; variable region.
OS      Homo sapiens.
XX
XX      EP325262-A.
PN
XX      26-JUL-1989.
PD
XX      20-JAN-1989; 89EP-00100913.
PF
XX      22-JAN-1988; 88US-00147351.
PR
XX      (GENO ) GEN HOSPITAL CORP.
PA
XX
XX      Seed B;
PI      WPI; 1989-214472/30.
XX      DR      N-PSDB; AAN90356.
XX
XX      Immunoglobulin-CD4 fusion proteins - used for treating HIV or SIV
XX      infections or detecting HIV or SIV in sample.
PT
XX
XX      Example; Table 1, Page 12-23; 68pp; English.
PS
XX
XX      The fusion protein genes of the invention pref. comprise cDNA sequences
XX      which encode CD4 or a fragment which binds gp120 ligated to an expression
XX      plasmid which encodes an antibody in which the variable region of the
XX      gene has been deleted (see WO87-02671). The CD4 portion of the fusion
XX      protein may comprise the complete CD4 sequence, the 370 AA extracellular
XX      region and the membrane spanning domain, or the extracellular region. The
XX      Ig heavy chain is pref. from IgM, IgG1 or IgG3. The following are
XX      specifically claimed: fusion proteins CD4H-gamma-1, CD4mu, CD4Pmu, CD4E-
XX      gamma1, and CD4mu (No. 67608), PCDAH-gamma (No. 67609) and PCDAE-gamma-1
XX      (No. 67610). The plasmid containing (PCDAH-gamma-1) has been deposited in
XX      E. coli (MC1061/F3) at the ATCC under accession number 67611. (Updated on
XX      25-MAR-2003 to correct PA field.)
SQ
XX
XX      Sequence 729 AA;
Query Match 93.0%; Score 3176.5; DB 1; Length 729;
Best Local Similarity 85.3%; Pred. No. 2.5e-163;
Matches 621; Conservative 0; Mismatches 4; Indels 103; Gaps 3;

```

```

Db      181  TWCTCTVQNQKVEFKIDIVLAFQKASSIVYKKEGGEVRSFPLAFTVEKLTSGELMW 240
|||||
Db      241  QAERASSSSKSWITTDLNKKEVSVKRVTDPRKLGKPLHLTLPLALPOYAGSGNLTLA 300
|||||
Db      241  QAERASSSSKSWITTDLNKKEVSVKRVTDPRKLGKPLHLTLPLALPOYAGSGNLTLA 300
|||||
Db      301  LEAKTKLHDEVNLVWMAATOLQNLTCGVGPTSPKLMSLKLENKAKYSKREKPVW 360
|||||
Db      301  LEAKTKLHDEVNLVWMAATOLQNLTCGVGPTSPKLMSLKLENKAKYSKREKPVW 360
|||||
Db      361  LNPEAGMWQCCLSDSGVLTLESNIKVLPTWSTPYPC----- 396
|||||
Db      361  LNPEAGMWQCCLSDSGVLTLESNIKVLPTWSTPYPC----- 396
|||||
Db      397  -----PAP----- 399
|||||
Db      421  TAAIGCLVKDYFPEPVTVSWNSGALTSGVHFPVAVLQSSGLYSLSVTVPSLSIGTQY 480
|||||
Db      400  -----EPKSCDKTHHC-----PELIGSPVFLPPPKDTLMISRTPE 437
|||||
Db      481  ICNVNHRPSNTKVDKKEVPSCDKTHHCPCCPABELIGSPVFLPPPKDTLMISRTPE 540
|||||
Db      438  VTCVAVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKE 497
|||||
Db      541  VTCVAVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDLNKE 600
|||||
Db      498  YKCVSNKALPAPIEKTISKAKGQPREPOVYTLPPSDELTKNQVSLTCLVKGFYPSDIA 557
|||||
Db      601  YKCVSNKALPAPIEKTISKAKGQPREPOVYTLPPSDELTKNQVSLTCLVKGFYPSDIA 660
|||||
Db      558  VEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVHNEALHNHYTQ 617
|||||
Db      661  VEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRMQQGNVFSCSVHNEALHNHYTQ 720
|||||
Db      618  KSLSLSPG 625
|||||
Db      721  KSLSLSPG 728
|||||
RESULT 6
AAB19507
ID      AAB19507 standard; protein; 729 AA.
XX
XX      AAB19507;
AC
XX      09-JAN-2001 (first entry)
DT
XX
DE      CD4-IgG1 fusion protein CH4Hgamma1.
XX
XX      CD4; IgG1; human; CD4Hgamma1; fusion protein; immunoglobulin; HIV; SIV;
XX      gp120; therapy; diagnosis.
XX
XX      Homo sapiens.
OS
XX
XX      Key
FH      Location/Qualifiers
FT      Protein
FT      1..395
FT      /note= "CD4 extracellular region"
FT      Protein
FT      400..729
FT      /note= "IgG1 heavy chain"
XX
XX      US6117656-A.
XX
XX      12-SEP-2000.
XX
XX      07-JUN-1995; 95US-00479353.
XX
XX      22-JAN-1988; 88US-00147351.
XX      23-JAN-1989; 89US-00299596.
XX      09-JUN-1992; 92US-00896781.
XX      12-APR-1993; 93US-00057952.
XX      04-FEB-1994; 94US-00191708.
XX

```

PA (GEHO) GEN HOSPITAL CORP.
 XX Seed B;
 PI
 DR WPI; 2000-586558/55.
 DR N-PSDB; AAA50660.
 XX
 PT CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or
 PT SIV.
 XX
 PS Example 1; Col 13-30; 39pp; English.
 XX
 CC The present sequence is that of fusion protein CD4Hgmam1 comprising the
 CC extracellular portion of CD4, which binds to HIV gp120, linked at its C-
 CC terminus to the human IgG1 heavy chain. To obtain the fusion protein, DNA
 CC encoding CD4 was linked to IgG1 DNA at the Hind3 site upstream of the CH1
 CC region (see AAA50660). Fusion protein CD4Hgmam1 and a nucleic acid
 CC encoding it are claimed. Also claimed are a vector comprising the nucleic
 CC acid, and a method of producing the fusion protein in secreted form using
 CC a transformed host cell. The fusion protein may further comprise a
 CC therapeutic agent, radiolabel or NMR imaging agent. The fusion protein
 CC can be administered to an animal (including humans) for treatment of HIV
 CC or SIV infection, and can also be used in assays for HIV or SIV, imaging
 CC and tissue stains. IgG1 fusion proteins such as CD4Hgmam1 provide both
 CC complement-mediated and cell-mediated immunity
 XX
 SQ Sequence 729 AA;
 Query Match 93.0%; Score 3176.5; DB 3; Length 729;
 Best Local Similarity 85.3%; Pred. No. 2.5e-163;
 Matches 621; Conservative 0; Mismatches 4; Indels 103; Gaps 3;
 QY 1 MNRGVPFRHLVLVQLALLPAATQGNKVLGKKGDVVELTCTASOKKSIOFHMKNNSQIK 60
 DB 1 MNRGVPFRHLVLVQLALLPAATQGNKVLGKKGDVVELTCTASOKKSIOFHMKNNSQIK 60
 QY 61 ILNGQGSFLTKGPKSLNDRADRSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEVOL 120
 DB 61 ILNGQGSFLTKGPKSLNDRADRSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEVOL 120
 QY 121 LVFGLTANSDTHLLOQSLTLTLSPGSSPSVQCRSPRKNIOGCKTLVSQLELDQSG 180
 DB 121 LVFGLTANSDTHLLOQSLTLTLSPGSSPSVQCRSPRKNIOGCKTLVSQLELDQSG 180
 QY 181 TWTCVTYQONQKVEFKIDIVLAFQKASSIYKKKEQVEPSFPFLATVEKLTGSGELMW 240
 DB 181 TWTCVTYQONQKVEFKIDIVLAFQKASSIYKKKEQVEPSFPFLATVEKLTGSGELMW 240
 QY 241 QAEBSASSKSWITFDLKNKEVSVRKVTQDPKLOMGKLLPHLTLPOALPOYAGSGNLTLA 300
 DB 241 QAEBSASSKSWITFDLKNKEVSVRKVTQDPKLOMGKLLPHLTLPOALPOYAGSGNLTLA 300
 QY 301 LEAKTGLHQBVLNVMKATQLOKNLTCEVWGPTSPKXLMSTLKLLENKAVSKREKPVWV 360
 DB 301 LEAKTGLHQBVLNVMKATQLOKNLTCEVWGPTSPKXLMSTLKLLENKAVSKREKPVWV 360
 QY 361 LNPBAGMWQCLISGQVLLLESNIKVLPTWSTPVPADPEASTKGPVFPPLAPSGSKTSGG 420
 DB 361 LNPBAGMWQCLISGQVLLLESNIKVLPTWSTPVPADPEASTKGPVFPPLAPSGSKTSGG 420
 QY 397 -----PAP----- 399
 DB 421 TAAIGCLVSTFPEPVTVMNSGALITSGVHTFPVAVLOSGLYSLSSVTVTPSSSLGTQTY 480
 QY 400 -----EPKSCDKTHTC-----PELLGSPVFLFPKPKDTLMISTRTPE 437
 DB 481 ICNVNHNKPSMTKVDKVKPEKSCDKTHTCPCPAPPELLGSPVFLFPKPKDTLMISTRTPE 540
 QY 438 VTCVAVDVSHEDPEVKKNMYVDGVEVFNNAKTKPREEOYNSYRVVSVLTVLHODMLNGKE 497
 DB 541 VTCVAVDVSHEDPEVKKNMYVDGVEVFNNAKTKPREEOYNSYRVVSVLTVLHODMLNGKE 600
 QY 498 YKCVSNKALPAPIEKTISKAKGQPRPEQVYTLTPPSBDELTKQNVSLTCLVKGPYPSDIA 557

DB 601 YKCVSNKALPAPIEKTISKAKGQPRPEQVYTLTPPSBDELTKQNVSLTCLVKGPYPSDIA 660
 QY 558 VEMESNGQENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGVNFCGVMEALHNHYTQ 617
 DB 661 VEMESNGQENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGVNFCGVMEALHNHYTQ 720
 QY 618 KSLSLSPG 625
 DB 721 KSLSLSPG 728
 RESULT 7
 AA59168
 ID AA59168 standard; protein; 729 AA.
 AC AA59168;
 XX 14-MAR-2000 (first entry)
 DT
 DE CD4-Ig fusion protein CD4Hgmam1.
 XX
 XX HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;
 KM secreted protein; SIV infection; medicament.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN CA1340741-C.
 PD 14-SEP-1999.
 XX
 PF 20-JAN-1989; 89CA-00588749.
 XX
 PR 20-JAN-1989; 89CA-00588749.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX Seed B;
 PI
 DR WPI; 2000-063015/06.
 DR N-PSDB; AA4248201.
 XX
 PT New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in
 PT the treatment of HIV or simian immunodeficiency virus infections.
 XX
 PS Example 1; Page 25-36; 89pp; English.
 XX
 CC The invention provides a fusion gene encoding a fusion protein that
 CC comprises an extracellular CD4 DNA sequence or its fragment which binds
 CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA
 CC sequence of an Ig heavy or light chain, where the DNA sequence encoding
 CC the variable region has been replaced with the DNA sequence which encodes
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is
 CC capable of being secreted. The fusion proteins are useful for treating
 CC HIV or SIV infections in animals, preferably humans. They are also useful
 CC for producing medicaments which can be used for treating HIV or SIV
 CC infections in humans. The present sequence represents the fusion protein
 CC CD4Hgmam1 where the CD4 is linked to human IgG1 at the Hind3 site
 CC upstream of the CH1 region
 XX
 SQ Sequence 729 AA;
 Query Match 93.0%; Score 3176.5; DB 3; Length 729;
 Best Local Similarity 85.3%; Pred. No. 2.5e-163;
 Matches 621; Conservative 0; Mismatches 4; Indels 103; Gaps 3;
 QY 1 MNRGVPFRHLVLVQLALLPAATQGNKVLGKKGDVVELTCTASOKKSIOFHMKNNSQIK 60
 DB 1 MNRGVPFRHLVLVQLALLPAATQGNKVLGKKGDVVELTCTASOKKSIOFHMKNNSQIK 60
 QY 61 ILNGQGSFLTKGPKSLNDRADRSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEVOL 120

```

Db      61 ILGNQSFLLTKGPSKLANDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEYQL 120
Qy      121 LVFGLTANSDTHLLQGSLLTLTLSPGSSPSVOCRSRGNIOGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLQGSLLTLTLSPGSSPSVOCRSRGNIOGKTLVSQLELDQSG 180
Qy      181 TWCTVLQONQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFTVEKLTGSGELMW 240
Db      181 TWCTVLQONQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFTVEKLTGSGELMW 240
Qy      241 QAEARASSKSWITFDLKNKEVSVKRVYODPKLQNGKKLPLHLTLPLQALPOYAGSGNLTLA 300
Db      241 QAEARASSKSWITFDLKNKEVSVKRVYODPKLQNGKKLPLHLTLPLQALPOYAGSGNLTLA 300
Qy      301 LEAKTGKHOEVNLVWRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKKEKPVWV 360
Db      301 LEAKTGKHOEVNLVWRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKKEKPVWV 360
Qy      361 LNPEAGMWQCILSDSGQVLLSNIKVLPTWSTPVPC----- 396
Db      361 LNPEAGMWQCILSDSGQVLLSNIKVLPTWSTPVHADPEASTKGSVFPPLAPSSKSTSGG 420
Qy      397 -----PAP----- 399
Db      421 TAAIGCLVKDYFPEPVTVSNMNSGALTSQVHTFPAVLQSSGLYSLSVTVTPSSSLGTQTY 480
Qy      400 -----EPKSCDKTHTC-----PELLGSPSVFLPPPKPDTLMISRTPE 437
Db      481 ICNVNHRKPSNTKVDKKEVPCSDTKHTPCPCPAPBELGSPSVFLPPPKPDTLMISRTPE 540
Qy      438 VTCVVDVSHDEPEVKFNMYVDGVEVHNATKPREEQNSTYRVVSVLTVLHODMLNGKE 497
Db      541 VTCVVDVSHDEPEVKFNMYVDGVEVHNATKPREEQNSTYRVVSVLTVLHODMLNGKE 600
Qy      498 YKCVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDELTKNOVSLTCLVKGYFSPDIA 557
Db      601 YKCVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDELTKNOVSLTCLVKGYFSPDIA 660
Qy      558 VEMESNQGPENNYKTPPVLDSDGSFLYSKLTVDKSRWQGNVFSQSVHMEALHNHYTQ 617
Db      661 VEMESNQGPENNYKTPPVLDSDGSFLYSKLTVDKSRWQGNVFSQSVHMEALHNHYTQ 720
Qy      618 KSLSLSPG 625
Db      721 KSLSLSPG 728

```

RESULT 8
ID AAYS1078 standard; protein, 729 AA.

```

XX AC AAYS1078;
XX DT 23-MAR-2000 (first entry)
XX DE Human fusion protein CD4H-1.
XX KM Fusion protein; human, CD4, IgG1, immunoglobulin, gp120;
XX KM anti-human immunodeficiency virus, CD4H-1.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN US6004781-A.
XX PD 21-DEC-1999.
XX PF 04-FEB-1994; 94US-00191708.
XX PR 22-JAN-1988; 88US-00147351.
XX PR 23-JAN-1989; 89US-00299596.
XX PR 09-JUN-1992; 92US-00896781.
XX PR 12-APR-1993; 93US-00057952.

```

```

XX PA (GENO ) GEN HOSPITAL CORP.
XX PI Seed B;
XX DR WPI; 2000-085792/07.
XX DR N-PSDB; AA244061.
XX PT Fusion protein useful for the treatment of human immunodeficiency virus.
XX PS Example 1; Col 15-30; 39pp; English.
XX CC This invention describes a novel nucleic acid (I) encoding a fusion
CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
CC chain (III). The products of the invention have anti-human
CC immunodeficiency virus (HIV) activity and are capable of binding to
CC gp120. The fusion protein is useful for treating human immunodeficiency
CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence
CC represents the fusion protein CD4H-1 which is constructed from CD4 linked
CC to human IgG1 upstream of the CH1 region
XX SQ Sequence 729 AA;
XX Qy Query Match 92.8%; Score 3166.5; DB 3; Length 729;
XX Qy Best Local Similarity 85.2%; Pred. No. 8.8e-163;
XX Qy Matches 620; Conservative 0; Mismatches 5; Indels 103; Gaps 3;
XX Qy 1 MNRGVPFRHLLVLTQALLPRAATQGNRVVLGKKDTELTCTASQKKSIOFHMNSNQIK 60
XX Qy 1 MNRGVPFRHLLVLTQALLPRAATQGNRVVLGKKDTELTCTASQKKSIOFHMNSNQIK 60
XX Qy 61 ILGNQSFLLTKGPSKLANDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEYQL 120
XX Qy 61 ILGNQSFLLTKGPSKLANDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEYQL 120
XX Qy 61 ILGNQSFLLTKGPSKLANDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEYQL 120
XX Qy 121 LVFGLTANSDTHLLQGSLLTLTLSPGSSPSVOCRSRGNIOGKTLVSQLELDQSG 180
XX Qy 121 LVFGLTANSDTHLLQGSLLTLTLSPGSSPSVOCRSRGNIOGKTLVSQLELDQSG 180
XX Qy 121 LVFGLTANSDTHLLQGSLLTLTLSPGSSPSVOCRSRGNIOGKTLVSQLELDQSG 180
XX Qy 181 TWCTVLQONQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFTVEKLTGSGELMW 240
XX Qy 181 TWCTVLQONQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFTVEKLTGSGELMW 240
XX Qy 241 QAEARASSKSWITFDLKNKEVSVKRVYODPKLQNGKKLPLHLTLPLQALPOYAGSGNLTLA 300
XX Qy 241 QAEARASSKSWITFDLKNKEVSVKRVYODPKLQNGKKLPLHLTLPLQALPOYAGSGNLTLA 300
XX Qy 241 QAEARASSKSWITFDLKNKEVSVKRVYODPKLQNGKKLPLHLTLPLQALPOYAGSGNLTLA 300
XX Qy 301 LEAKTGKHOEVNLVWRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKKEKPVWV 360
XX Qy 301 LEAKTGKHOEVNLVWRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSKKEKPVWV 360
XX Qy 361 LNPEAGMWQCILSDSGQVLLSNIKVLPTWSTPVPC----- 396
XX Qy 361 LNPEAGMWQCILSDSGQVLLSNIKVLPTWSTPVPC----- 396
XX Qy 361 LNPEAGMWQCILSDSGQVLLSNIKVLPTWSTPVPC----- 396
XX Qy 397 -----PAP----- 399
XX Qy 421 TAAIGCLVKDYFPEPVTVSNMNSGALTSQVHTFPAVLQSSGLYSLSVTVTPSSSLGTQTY 480
XX Qy 400 -----EPKSCDKTHTC-----PELLGSPSVFLPPPKPDTLMISRTPE 437
XX Qy 481 ICNVNHRKPSNTKVDKKEVPCSDTKHTPCPCPAPBELGSPSVFLPPPKPDTLMISRTPE 540
XX Qy 438 VTCVVDVSHDEPEVKFNMYVDGVEVHNATKPREEQNSTYRVVSVLTVLHODMLNGKE 497
XX Qy 541 VTCVVDVSHDEPEVKFNMYVDGVEVHNATKPREEQNSTYRVVSVLTVLHODMLNGKE 600
XX Qy 498 YKCVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDELTKNOVSLTCLVKGYFSPDIA 557
XX Qy 601 YKCVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDELTKNOVSLTCLVKGYFSPDIA 660
XX Qy 558 VEMESNQGPENNYKTPPVLDSDGSFLYSKLTVDKSRWQGNVFSQSVHMEALHNHYTQ 617

```

Db 661 VEMESNGQPENNKTPPVLDSDGSFPLYSKLTVDSRMQGNVFCSSVYHEALHNHYQ 720

QY 618 KSLSLSPG 625

Db 721 KSLSLSPG 728

RESULT 9

AA51082

ID AAY51082 standard; protein; 616 AA.

XX AAY51082;

AC 23-MAR-2000 (first entry)

XX Human fusion protein CD4Bgamal.

XX Fusion protein; human; CD4; IgG1; immunoglobulin; gp120;

XX anti-human immunodeficiency virus; CD4Bgamal.

XX Homo sapiens.

XX Synthetic.

XX US6004781-A.

XX 21-DEC-1999.

XX 04-FEB-1994; 94US-00191708.

XX 22-JAN-1988; 88US-00147351.

XX 23-JAN-1989; 89US-00299596.

XX 09-JUN-1992; 92US-00896781.

XX 12-APR-1993; 93US-00057952.

XX (GEHO) GEN HOSPITAL CORP.

XX Seed B;

XX WPI; 2000-085792/07.

XX N-PSDB; AAZ44065.

XX Fusion protein useful for the treatment of human immunodeficiency virus.

XX Example 1; Col 59-70; 39pp; English.

XX This invention describes a novel nucleic acid (I) encoding a fusion

XX protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)

XX and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light

XX chain (III). The products of the invention have anti-human

XX immunodeficiency virus (HIV) activity and are capable of binding to

XX gp120. The fusion protein is useful for treating human immunodeficiency

XX virus (HIV) or simian immunodeficiency virus (SIV). This sequence

XX represents the fusion protein CD4Bgamal which is constructed from CD4

XX linked to human IgG1 upstream of the hinge region

XX Sequence 616 AA;

Query Match 92.6%; Score 3163; DB 3; Length 616;

Best Local Similarity 97.6%; Pred. No. 1,1e-162;

Matches 610; Conservative 0; Mismatches 5; Indels 10; Gaps 2;

QY 1 MARGVPRRHLLVQLLLPAATGKAVVGLKKGDYELTCTASQKKSIOFHKNSQIK 60

Db 1 MARGVPRRHLLVQLLLPAATGKAVVGLKKGDYELTCTASQKKSIOFHKNSQIK 60

QY 61 ILGNQSFLLTKGPKLNDRAISRSLMDQGNFLLIINGLKIENSDTYICEVEDQKEPVQL 120

Db 61 ILGNQSFLLTKGPKLNDRAISRSLMDQGNFLLIINGLKIENSDTYICEVEDQKEPVQL 120

QY 121 LVFGLTANSDTHLLDQGSLLTLTSPGSSPSVQCRSPKNTQGGKTLISVQLDQSG 180

Db 121 LVFGLTANSDTHLLDQGSLLTLTSPGSSPSVQCRSPKNTQGGKTLISVQLDQSG 180

QY 181 TWCTVLOQKKEFKIDIVLAFQKASSIYKKEGEVFEFPLAFVTEKLTGSGELMW 240

Db 181 TWCTVLOQKKEFKIDIVLAFQKASSIYKKEGEVFEFPLAFVTEKLTGSGELMW 240

QY 241 QAERASSSKSWITFDLKNKEVSVKVETODPKLQMGKULPLHLTLPOALPOYAGSGNLTLA 300

Db 241 QAERASSSKSWITFDLKNKEVSVKVETODPKLQMGKULPLHLTLPOALPOYAGSGNLTLA 300

QY 301 LEAKTGKLEHVEVNLVYMRATQLOKULTCVWGPSTSPKMLSLKLENKAKVSKREPVWV 360

Db 301 LEAKTGKLEHVEVNLVYMRATQLOKULTCVWGPSTSPKMLSLKLENKAKVSKREPVWV 360

QY 361 LNPEAGMOCCLSDSGQVLEESNIKYLPWTSFVPCPAPRPSCDKTHCPPELLGSPSYF 420

Db 361 LNPEAGMOCCLSDSGQVLEESNIKYLPWTSFVPCPAPRPSCDKTHCPPELLGSPSYF 420

QY 421 LEPPKPKDLMISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTRPREQYNSTYR 480

Db 421 LEPPKPKDLMISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTRPREQYNSTYR 480

QY 441 LEPPKPKDLMISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTRPREQYNSTYR 470

Db 441 LEPPKPKDLMISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTRPREQYNSTYR 470

QY 481 VVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKX 540

Db 481 VVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKX 540

QY 471 VVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKX 530

Db 471 VVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKX 530

QY 541 QVSLTCLVKGFPSPDI AVEESNGQPENNKTPPVLDSDGSFPLYSKLTVDSRMQGN 600

Db 541 QVSLTCLVKGFPSPDI AVEESNGQPENNKTPPVLDSDGSFPLYSKLTVDSRMQGN 600

QY 601 VFSCSYMEALHNHYTQKSLSPG 625

Db 601 VFSCSYMEALHNHYTQKSLSPG 625

QY 591 VFSCSYMEALHNHYTQKSLSPG 615

Db 591 VFSCSYMEALHNHYTQKSLSPG 615

RESULT 10

AA59172

ID AAY59172 standard; protein; 616 AA.

XX AAY59172;

XX 14-MAR-2000 (first entry)

XX CD4-Ig fusion protein CD4Bgamal.

XX HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;

XX secreted protein; SIV infection; medicament.

XX Synthetic.

XX Homo sapiens.

XX CA1340741-C.

XX 14-SEP-1999.

XX 20-JAN-1989; 89CA-00588749.

XX 20-JAN-1989; 89CA-00588749.

XX (GEHO) GEN HOSPITAL CORP.

XX Seed B;

XX WPI; 2000-063015/06.

XX N-PSDB; AAZ48205.

XX New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in

XX the treatment of HIV or simian immunodeficiency virus infections.

XX Example 1; Page 61-68; 89pp; English.

XX The invention provides a fusion gene encoding a fusion protein that

XX comprises an extracellular CD4 DNA sequence or its fragment which binds

XX to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA

CC sequence of an Ig heavy or light chain, where the DNA sequence encoding
 CC the variable region has been replaced with the DNA sequence which encodes
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is
 CC capable of being secreted. The fusion proteins are useful for treating
 CC HIV or SIV infections in animals, preferably humans. They are also useful
 CC for producing medicaments which can be used for treating HIV or SIV
 CC infections in humans. The present sequence represents the fusion protein
 CC CD4Bgamma where the CD4 is linked to human IgG1 at the BamI site
 CC downstream from the hinge region

XX Sequence 616 AA:

Query Match 92.6%; Score 3163; DB 3; Length 616;

Best Local Similarity 97.6%; Pred. No. 1,1e-162;

Matches 610; Conservative 0; Mismatches 5; Indels 10; Gaps 2;

QY 1 MNRGVPRHLLLVQLALPAAQGNKRVLGKGGDTVELTCTASOKKSIOFHMKNNOIK 60
 DB 1 MNRGVPRHLLLVQLALPAAQGNKRVLGKGGDTVELTCTASOKKSIOFHMKNNOIK 60
 QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOL 120
 DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOL 120
 QY 121 LVFGLTANSDTHLQGGSLTLTLSPSSPSVOCSPRGKNIQGGKTLVSQLELDQSG 180
 DB 121 LVFGLTANSDTHLQGGSLTLTLSPSSPSVOCSPRGKNIQGGKTLVSQLELDQSG 180
 QY 181 TWCTCTVLONQKKEFKIDIVLAFQKASSIYKKEGEVFSPLATVEKLTGSGELMW 240
 DB 181 TWCTCTVLONQKKEFKIDIVLAFQKASSIYKKEGEVFSPLATVEKLTGSGELMW 240
 QY 241 QAERASSSKSMITFDLKNKEVSVKRVYODPKLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
 DB 241 QAERASSSKSMITFDLKNKEVSVKRVYODPKLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
 QY 301 LEAKTGKLGHOEVNLVVRATOLQKNLTCEVWGPTSPKMLSLKLENKAVSKREKPYWV 360
 DB 301 LEAKTGKLGHOEVNLVVRATOLQKNLTCEVWGPTSPKMLSLKLENKAVSKREKPYWV 360
 QY 361 LNPBAGMWQCLLSDSGQVLLSNNIKVLPWTSTPVPCAPBPKSCDKHTTCBELLGSPVF 420
 DB 361 LNPBAGMWQCLLSDSGQVLLSNNIKVLPWTSTPVPCAPBPKSCDKHTTCBELLGSPVF 420
 QY 421 LFPKPKDPTLMISTREPTCVVVDVSHDEDEKFNWYVDGEVHNATKPREEOYNSTYR 480
 DB 421 LFPKPKDPTLMISTREPTCVVVDVSHDEDEKFNWYVDGEVHNATKPREEOYNSTYR 480
 QY 481 VVSVLTVLHODMLNGKEYCKVSNKALPAPIKTISSKAGOPREPOVYTLPPSRDELTKN 540
 DB 481 VVSVLTVLHODMLNGKEYCKVSNKALPAPIKTISSKAGOPREPOVYTLPPSRDELTKN 540
 QY 541 QVSLTCLVKGFPSPDIAVEMESNQPENNYKTPRPVLSDSGSFLYSKLTVDKSRMOQGN 600
 DB 541 QVSLTCLVKGFPSPDIAVEMESNQPENNYKTPRPVLSDSGSFLYSKLTVDKSRMOQGN 600
 QY 601 VFSCSVHGEALHNHYTQKSLSLSPG 625
 DB 601 VFSCSVHGEALHNHYTQKSLSLSPG 625
 QY 591 VFSCSVHGEALHNHYTQKSLSLSPG 615
 DB 591 VFSCSVHGEALHNHYTQKSLSLSPG 615

RESULT 11

ID AAP93012 standard; protein; 614 AA.

XX AAP93012;

XX 25-MAR-2003 (revised)

DT 03-AUG-1992 (first entry)

XX Genetic construct which encodes CD4 linked to human IgG1 at the BamI site
 DE downstream from the hinge region (fusion protein CD4Bgamma).

KW Fusion protein; immunoglobulin-like molecule; HIV, SIV, therapy;
 KM diagnosis; CD4; gp120; binding fragment; glycoprotein; variable region.

OS Homo sapiens.

PN EPJ25262-A.

XX 26-JUL-1989.

PF 20-JAN-1989; 89EP-00100913.

PR 22-JAN-1988; 88US-00147351.

PA (GENO) GEN HOSPITAL CORP.

XX Seed B;

DR WP1; 1989-214472/30.

DR N-PSDB; AAN90360.

PT Immunoglobulin-CD4 fusion proteins - used for treating HIV or SIV
 PT infections or detecting HIV or SIV in sample.

PS Example; Table 5, Page 48-55; 68pp; English.

CC The fusion protein genes of the invention pref. comprises cDNA sequences
 CC which encode CD4 or a fragment which binds gp120 ligated to an expression
 CC plasmid which encodes an antibody in which the variable region of the
 CC gene has been deleted (see WO87-02671). The CD4 portion of the fusion
 CC protein may comprise the complete CD4 sequence, the 370 AA extracellular
 CC region and the membrane spanning domain, or the extracellular region. The
 CC Ig heavy chain is pref. from IgM, IgG1 or IgG3. The following are
 CC specifically claimed: fusion proteins CD4Bgamma, CD4Bmu, CD4Bnu,
 CC CD4Bdelta, and CD4Bmu (No. 67608), pCD4Bgamma (No. 67609) and
 CC pCD4Bdelta (No. 67610). (Updated on 25-MAR-2003 to correct PA field.)

SO Sequence 614 AA:

Query Match 91.7%; Score 3132; DB 1; Length 614;

Best Local Similarity 97.0%; Pred. No. 5.3e-161;

Matches 606; Conservative 0; Mismatches 7; Indels 12; Gaps 3;

QY 1 MNRGVPRHLLLVQLALPAAQGNKRVLGKGGDTVELTCTASOKKSIOFHMKNNOIK 60
 DB 1 MNRGVPRHLLLVQLALPAAQGNKRVLGKGGDTVELTCTASOKKSIOFHMKNNOIK 60
 QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOL 120
 DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOL 120
 QY 121 LVFGLTANSDTHLQGGSLTLTLSPSSPSVOCSPRGKNIQGGKTLVSQLELDQSG 180
 DB 121 LVFGLTANSDTHLQGGSLTLTLSPSSPSVOCSPRGKNIQGGKTLVSQLELDQSG 180
 QY 181 TWCTCTVLONQKKEFKIDIVLAFQKASSIYKKEGEVFSPLATVEKLTGSGELMW 240
 DB 181 TWCTCTVLONQKKEFKIDIVLAFQKASSIYKKEGEVFSPLATVEKLTGSGELMW 240
 QY 241 QAERASSSKSMITFDLKNKEVSVKRVYODPKLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
 DB 241 QAERASSSKSMITFDLKNKEVSVKRVYODPKLQMGKPLHLTLPLQALPOYAGSGNLTIA 300
 QY 301 LEAKTGKLGHOEVNLVVRATOLQKNLTCEVWGPTSPKMLSLKLENKAVSKREKPYWV 360
 DB 301 LEAKTGKLGHOEVNLVVRATOLQKNLTCEVWGPTSPKMLSLKLENKAVSKREKPYWV 360
 QY 361 LNPBAGMWQCLLSDSGQVLLSNNIKVLPWTSTPVPCAPBPKSCDKHTTCBELLGSPVF 420
 DB 361 LNPBAGMWQCLLSDSGQVLLSNNIKVLPWTSTPVPCAPBPKSCDKHTTCBELLGSPVF 420
 QY 421 LFPKPKDPTLMISTREPTCVVVDVSHDEDEKFNWYVDGEVHNATKPREEOYNSTYR 480
 DB 421 LFPKPKDPTLMISTREPTCVVVDVSHDEDEKFNWYVDGEVHNATKPREEOYNSTYR 480
 QY 411 LFPKPKDPTLMISTREPTCVVVDVSHDEDEKFNWYVDGEVHNATKPREEOYNSTYR 470
 DB 411 LFPKPKDPTLMISTREPTCVVVDVSHDEDEKFNWYVDGEVHNATKPREEOYNSTYR 470

```

QY 481 VSVLTVLVHODMNGEKYCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKN 540
DB 471 VSVLTVLVHODMNGEKYCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKN 528
QY 541 QVSLTCLVKGFPSPDIADVEMESNGOPENNKTTPPVLDSDGSFFLYSKLTVDSKSRMOQGN 600
DB 529 QVSLTCLVKGFPSPDIADVEMESNGOPENNKTTPPVLDSDGSFFLYSKLTVDSKSRMOQGN 588
QY 601 VFSCSVMEHALHNHYTKSLSLSPG 625
DB 589 VFSCSVMEHALHNHYTKSLSLSPG 613

RESULT 12
AAB19511
ID AAB19511 standard; protein; 616 AA.
XX
AC AAB19511;
XX
DT 09-JAN-2001 (first entry)
XX
DE CD4-IgG1 fusion protein CH4Bgammal.
XX
KM CD4; IgG1; human; CD4Bgammal; fusion protein; immunoglobulin; HIV; SIV;
KM gp120; therapy; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..395
FT Protein /note="CD4 extracellular region"
FT Protein 400..616
FT Protein /note="IgG1 heavy chain"
XX
XX US6117656-A.
XX
PD 12-SEP-2000.
XX
PF 07-JUN-1995; 95US-00479353.
XX
PR 22-JAN-1988; 88US-00147351.
PR 23-JAN-1989; 89US-00299596.
PR 09-JUN-1992; 92US-0086781.
PR 12-APR-1993; 93US-00057952.
PR 04-FEB-1994; 94US-00191708.
XX
PA (GENE) GEN HOSPITAL CORP.
XX
PI Seed B;
XX
DR WPI; 2000-586558/55.
DR N-PSDB; AAA50664.
XX
PT CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or
PT SIV.
XX
PS Example 1; Col 59-70; 39pp; English.
XX
CC The present sequence is that of fusion protein CD4Bgammal comprising the
CC extracellular portion of CD4, which binds to HIV gp120, linked at its C-
CC terminus to the human IgG1 heavy chain. To obtain the fusion protein, DNA
CC encoding CD4 was linked to IgG1 DNA at the BamI site downstream of the
CC hinge region (see AAA50664). Fusion protein CD4Bgammal and a nucleic acid
CC encoding it are claimed. Also claimed are a vector comprising the nucleic
CC acid, and a method of producing the fusion protein in secreted form using
CC a transformed host cell. The fusion protein may further comprise a
CC therapeutic agent, radiolabel or NMR imaging agent. The fusion protein
CC can be administered to an animal (including humans) for treatment of HIV
CC or SIV infection, and can also be used in assays for HIV or SIV, imaging
CC and tissue stains. IgG1 fusion proteins such as CD4Bgammal provide both
CC complement-mediated and cell-mediated immunity
XX

```

```

SQ Sequence 616 AA;
Query Match 90.0%; Score 3071; DB 3; Length 616;
Best Local Similarity 96.0%; Pred. No. 1e-157;
Matches 599; Conservative 0; Mismatches 15; Indels 10; Gaps 2;

QY 1 MKRGVPRHLLVLTQALLPATQGNKYVIGKKGPVETLTASQKSIQFMKNSNQIK 60
DB 1 MKRGVPRHLLVLTQALLPATQGNKYVIGKKGPVETLTASQKSIQFMKNSNQIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFLLIKNLIKEDSPYICEVEDQKEEYVL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFLLIKNLIKEDSPYICEVEDQKEEYVL 120
QY 121 LVFGILTANSDTHLQGGSLTLTLESPPGSSPSVQCRSPKCKIQQCKTISVSQLELDSDG 180
DB 121 LVFGILTANSDTHLQGGSLTLTLESPPGSSPSVQCRSPKCKIQQCKTISVSQLELDSDG 180
QY 181 TWTCVTLQKQKVEFIDIVLAFOKASSIVYKKEGEQVEFAPLAFTVEKLTGSGELMW 240
DB 181 TWTCVTLQKQKVEFIDIVLAFOKASSIVYKKEGEQVEFAPLAFTVEKLTGSGELMW 240
QY 241 QAEKASSSKSWITFDLKNKEVSXKVTODPKLQMGKPLHLTLPOALPOYAGSGNLTLA 300
DB 241 QAEKASSSKSWITFDLKNKEVSXKVTODPKLQMGKPLHLTLPOALPOYAGSGNLTLA 300
QY 301 LEAKTGKLEHENVLYMVRATQKQNLTCGVMGPTSPKMLSLKLENKEAKVSKREKPVNV 360
DB 301 LEAKTGKLEHENVLYMVRATQKQNLTCGVMGPTSPKMLSLKLENKEAKVSKREKPVNV 360
QY 361 LNPEAGMOCCLSDSGQVLLAESNIKVLPTWSTPVPCEPAPEPSCDKTHTCPPELLGSPSYF 420
DB 361 LNPEAGMOCCLSDSGQVLLAESNIKVLPTWSTPVPCEPAPEPSCDKTHTCPPELLGSPSYF 420
QY 421 LEPPKPKDTLMTSRPEVTCVVVDVSHEDPEVKFMYVDGVENHNAKTPREBOINSTR 480
DB 421 LEPPKPKDTLMTSRPEVTCVVVDVSHEDPEVKFMYVDGVENHNAKTPREBOINSTR 480
QY 481 VSVLTVLVHODMNGEKYCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKN 540
DB 471 VSVLTVLVHODMNGEKYCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKN 530
QY 541 QVSLTCLVKGFPSPDIADVEMESNGOPENNKTTPPVLDSDGSFFLYSKLTVDSKSRMOQGN 600
DB 531 QVSLTCLVKGFPSPDIADVEMESNGOPENNKTTPPVLDSDGSFFLYSKLTVDSKSRMOQGN 590
QY 601 VFSCSVMEHALHNHYTKSLSLSP 624
DB 591 VFSCSVMEHALHNHYTKSLSLSP 614

RESULT 13
AAR26531
ID AAR26531 standard; protein; 534 AA.
XX
AC AAR26531;
XX
DT 25-MAR-2003 (revised)
DT 28-JAN-1993 (first entry)
XX
DE Sequence of CD4-IgG1 chimeric heavy chain heterotrimer.
XX
KM CD4-gamma 1 chimeric heavy chain homodimer; expression vector; HIV;
KM therapy; diagnostic agent; inhibition.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 205..302
FT Region /label=CH1
FT Region 303..317
FT Region /label=hinge
FT Region 318..427

```


Query Match 62.7%; Score 2140; DB 2; Length 435;
 Best Local Similarity 68.0%; Pred. No. 1.2e-107; Indels 200; Gaps 2;
 Matches 428; Conservative 0; Mismatches 1; Indels 200; Gaps 2;

QY 2 NRGVFFRHLILVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNSNQIKI 61
 DB 1 NRGVFFRHLILVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNSNQIKI 60

QY 62 LGNQGSLFKGSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 121
 DB 61 LGNQGSLFKGSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120

QY 122 VFGILTANSDTHLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGT 181
 DB 121 VFGILTANSDTHLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGT 180

QY 182 WTCITVLONOKKVEFKIDIVLAFOKASSIVYKKEGQEVESFLAFTVEKLTGSGELMW 241
 DB 181 WTCITVLONOKKVEFKIDIVLAFOKASSIVYKKEGQEVESFLAFTVEKLTGSGELMW 240

QY 242 AERASSSKSWITFDLKNKEVSKRVTDPKLQMGKLLPHLTLPOALPOYAGSGLTLAL 301
 DB 204 ----- 203

QY 302 EAKTGKLGHOEVNLVVMARATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPVWL 361
 DB 204 ----- 203

QY 362 NPEAGMWQCLLSDSGOVLLESNIKVLPTWSTPVPCEPAPEPKSCDKHTTC-----PELLGG 416
 DB 204 -----EPKSCDKHTTCPCPAPELLGG 225

QY 417 PSVFLFPKPKDTLMISTPEVTCVVVDVSHEDPEVKFNNVYDGVENNAKTKRREQYN 476
 DB 226 PSVFLFPKPKDTLMISTPEVTCVVVDVSHEDPEVKFNNVYDGVENNAKTKRREQYN 285

QY 477 STYVAVSLTTLHODMNGKYEKCKVSKNKLPAIEKTIISKAKQPREPOVYTLPSRDE 536
 DB 286 STYVAVSLTTLHODMNGKYEKCKVSKNKLPAIEKTIISKAKQPREPOVYTLPSRDE 345

QY 537 LTKNQSILTCLVKGFPSDIAVEMESNGQEPENNYKTPPVLDSDGSFFLYSKLTVDSRW 596
 DB 346 LTKNQSILTCLVKGFPSDIAVEMESNGQEPENNYKTPPVLDSDGSFFLYSKLTVDSRW 405

QY 597 QCGNVFSCSVMEHLAHNYTOKSLSPG 625
 DB 406 QCGNVFSCSVMEHLAHNYTOKSLSPG 434

RESULT 15
 AAR26783 standard; protein, 530 AA.
 AC AAR26783;
 AC 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 06-FEB-1993 (first entry)
 XX CD4-IgG2 chimeric heavy chain.
 XX
 KW homodimer; soluble CD4; T cell receptor; CD4 antigen; high recovery;
 KW chimeric; increased serum half life; HIV infection; AIDS; ss.
 XX
 OS Homo sapiens.
 OS Chimeric.
 XX
 FT Key location/Qualifiers
 FT Domain 1..205
 FT Domain /label= CD4 domain
 FT Domain 206..302
 FT Domain /label= CH1 domain
 FT Domain 303..312

FT Domain /label= hinge domain
 FT 313..423
 FT /label= CH2 domain
 FT Domain 424..530
 FT /label= CH3 domain

PN W09213947-A1.
 XX
 XX 20-ANG-1992.
 PD
 PF 10-FEB-1992; 92WO-US001143.
 XX
 PR 08-FEB-1991; 91US-00653684.
 PA (PROG-) PROGENICS PHARM INC.
 XX
 PI Beaudry GA, Madden PJ;
 XX
 DR WPI; 1992-300034/36.
 DR N-PSDB; AAQ28089.
 PT
 PT CD4-gamma-2 and CD4-IgG2 chimera(s) and expression vectors - for
 PS treatment, prevention and diagnosis of HIV infection.
 XX
 PS Claim 15; Fig 4; 90pp; English.

CC This sequence represents a CD4-IgG2 chimeric heavy chain heterotrimer
 CC it was produced by expression of the coding mutagenized cDNA (produced as
 CC described in AAQ28089) in Dhfr-CHO cells. The protein is efficiently
 CC assembled intracellularly and effectively secreted from mammalian cells
 CC pref. CHO, COS, or myeloma cells as a heterotrimer, enabling high
 CC recovery and purification from the medium of cells expressing it. It
 CC possesses increased serum half-life and has increased avidity for HIV cf.
 CC heavy chain dimers. It can inhibit HIV infection of CD4+ cells and block
 CC the spread of HIV infection within a patient. Attachment to a detectable
 CC marker makes it useful in an assay for HIV or SIV infection, and it can
 CC also be linked to toxins, eg diphtheria, Pseudomonas exotoxin A (domains
 CC I or II) or the deglycosylated A-chain of ricin. (Updated on 25-MAR-2003
 CC to correct PN field.) (Updated on 24-OCT-2003 to standardise OS field)

XX
 SQ Sequence 530 AA;
 QY
 DB
 QY 1 MNRGVFFRHLILVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNSNQIK 60
 DB 1 MNRGVFFRHLILVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNSNQIK 60

QY 61 ILNQGSLFKGSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120
 DB 61 ILNQGSLFKGSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120

QY 121 LVFGLTANSDTHLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGT 180
 DB 121 LVFGLTANSDTHLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGT 180

QY 181 TWCTITVLONOKKVEFKIDIVLAFOKASSIVYKKEGQEVESFLAFTVEKLTGSGELMW 240
 DB 181 TWCTITVLONOKKVEFKIDIVLAFOKASSIVYKKEGQEVESFLAFTVEKLTGSGELMW 240

QY 241 QAEARASSSKSWITFDLKNKEVSKRVTDPKLQMGKLLPHLTLPOALPOYAG---SGNL 297
 DB 241 QAEARASSSKSWITFDLKNKEVSKRVTDPKLQMGKLLPHLTLPOALPOYAG---SGNL 297

QY 298 TLALBAKTKLGHQEVNLVVMARATOL-OKNLTCEVWGPTSPKMLSLKENKAKVSKREK 356
 DB 298 TLALBAKTKLGHQEVNLVVMARATOL-OKNLTCEVWGPTSPKMLSLKENKAKVSKREK 356

QY 357 PVAVLNPEAGMWQCLLSDSGOVLLESNIKVLPTWSTPVP-----PCPAPEPKSCDKHTTC 411
 DB 293 P-----SNTKVDKTVKRCVCEPCPCPAP----- 317


```

XX Immunoglobulin; chelator; chimeric; HIV; human immunodeficiency virus.
XX Homo sapiens.
XX US6177549-B1.
XX 23-JUN-2001.
XX 10-JUN-1999; 99US-00329916.
XX 07-AUG-1992; 92US-00927931.
XX 06-AUG-1993; 93MO-US007422.
XX 03-FEB-1995; 95US-00379516.
XX 07-JUN-1995; 95US-00477460.
XX (PROG-) PROGENICS PHARM INC.
XX
XX Maddon PJ, Allaway GP;
XX
XX WPI; 2001-158582/16.
XX
XX Immunoglobulin for treating human immunodeficiency virus-infected
XX subject, consists of cytotoxic radionuclide linked to heterotetramer
XX comprising two chimeric CD4-IgG2 heavy chains and two chimeric CD4-
XX kappa light chains.
XX
XX Disclosure; Fig 4; 43pp; English.
XX
XX The present invention relates to an immunoglobulin, comprising a
XX cytotoxic radionuclide linked, directly or via a bifunctional chelator,
XX to a heterotetramer having two chimeric CD4-IgG2 heavy chains encoded by
XX an expression vector CD4-IgG2HC-PRCMV and two chimeric CD4-kappa light
XX chains encoded by an expression vector CD4-kLC-PRCMV. The invention is
XX useful for killing human immunodeficiency virus (HIV)-infected cells, for
XX the treatment and prevention of infection with HIV
XX
XX Sequence 530 AA;
XX
XX Query Match 62.0%; Score 2116; DB 4; Length 530;
XX Best Local Similarity 68.9%; Pred. No. 2.8e-106;
XX Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;
XX
XX 1 MNRGVPRRLHLLVQLALPAATQGNKRVLGKGDVVELTCTASOKKSIQFHKNSNQIK 60
XX 1 MNRGVPRRLHLLVQLALPAATQGNKRVLGKGDVVELTCTASOKKSIQFHKNSNQIK 60
XX
XX 61 ILGNQGSFLTKGSKINDRADSRSLMDQGNPFIITKNLKIETSDPTICVEEDQKEEVL 120
XX 61 ILGNQGSFLTKGSKINDRADSRSLMDQGNPFIITKNLKIETSDPTICVEEDQKEEVL 120
XX
XX 61 ILGNQGSFLTKGSKINDRADSRSLMDQGNPFIITKNLKIETSDPTICVEEDQKEEVL 120
XX 61 ILGNQGSFLTKGSKINDRADSRSLMDQGNPFIITKNLKIETSDPTICVEEDQKEEVL 120
XX
XX 121 LVFGLTANSDTHLLQGGSLTLLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
XX 121 LVFGLTANSDTHLLQGGSLTLLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
XX
XX 121 LVFGLTANSDTHLLQGGSLTLLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
XX 121 LVFGLTANSDTHLLQGGSLTLLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
XX
XX 181 TWCTCTVLQNKAKVEFKIDIVYLAFOKASSTIVYKKEGQVFSPLATVYKLTGSGELW 240
XX 181 TWCTCTVLQNKAKVEFKIDIVYLAFOKASSTIVYKKEGQVFSPLATVYKLTGSGELW 240
XX
XX 181 TWCTCTVLQNKAKVEFKIDIVYLAFOKASSTIVYKKEGQVFSPLATVYKLTGSGELW 240
XX 181 TWCTCTVLQNKAKVEFKIDIVYLAFOKASSTIVYKKEGQVFSPLATVYKLTGSGELW 240
XX
XX 241 QABRASSKSMWIFDLKNEKVSRYVQDPKLGKGLPHLTLPLQALPOYAG--SGNL 297
XX 241 QABRASSKSMWIFDLKNEKVSRYVQDPKLGKGLPHLTLPLQALPOYAG--SGNL 297
XX
XX 217 -----PCSRSTSESTALGCLVYKFPPEPVYVNSGALTSQVH 255
XX 217 -----PCSRSTSESTALGCLVYKFPPEPVYVNSGALTSQVH 255
XX
XX 298 TLALAKTGLHGVNLVWRATQL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
XX 298 TLALAKTGLHGVNLVWRATQL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
XX
XX 256 TPAVAVQSSGLYSLSVYVPSNFGQTYTCNV-----DHK 292
XX 256 TPAVAVQSSGLYSLSVYVPSNFGQTYTCNV-----DHK 292
XX
XX 357 PAVVAVLPEAGMOCCLSDSGVLLSNIKXLPWTSTV-----PCPAPPEKSCDKTHTCP 411
XX 357 PAVVAVLPEAGMOCCLSDSGVLLSNIKXLPWTSTV-----PCPAPPEKSCDKTHTCP 411
XX
XX 293 P-----SNTKVKTVRKCCVCEPCPAPP----- 317
XX 293 P-----SNTKVKTVRKCCVCEPCPAPP----- 317
XX
XX 412 ELTGSPVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGEVHNAAKTTPR 471
XX 412 ELTGSPVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGEVHNAAKTTPR 471
XX

```

```

DB 318 --VAGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGEVHNAAKTTPR 375
DB 472 EQQVNSTRYVSVLVVHODMUNGKYEYCKKSNKALPAPIETKISAKQPPREPOVYTP 531
DB 376 EQQVNSTRYVSVLVVHODMUNGKYEYCKKSNKALPAPIETKISAKQPPREPOVYTP 435
DB 532 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTIV 591
DB 436 PSREBMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTIV 495
DB 592 DKSRRQOGNVFSCSVMEHALHNHYTQKSLSLSPG 625
DB 496 DKSRRQOGNVFSCSVMEHALHNHYTQKSLSLSPG 529
DB
XX
XX RESULT 18
XX AAB80884
XX ID AAB80884 standard; protein; 530 AA.
XX
XX AAB80884;
XX
XX 29-MAY-2001 (first entry)
XX
XX Human CD4-IgG2 chimeric heavy chain.
XX
XX Human; Anti-HIV; CD4-IgG2 chimeric heterotetramer;
XX immunoglobulin gamma 2.
XX
XX Homo sapiens.
XX
XX US6187748-B1.
XX
XX 13-FEB-2001.
XX
XX 07-JUN-1995; 95US-00485372.
XX
XX 08-FEB-1991; 91US-00653684.
XX 10-FEB-1992; 92MO-US001143.
XX 08-DEC-1992; 92US-00960440.
XX
XX (PROG-) PROGENICS PHARM INC.
XX
XX Maddon PJ, Beaudry GA;
XX
XX WPI; 2001-264981/27.
XX N-PSDB; AAF77830.
XX
XX Inhibiting human immunodeficiency virus (HIV) infection of a CD4+ cell,
XX or treating a subject having CD4+ cells infected with HIV involves using
XX CD4-IgG2 chimeric heterotetramer to form a complex with the HIV.
XX
XX Disclosure; Fig 4; 55pp; English.
XX
XX The present invention relates to a method for inhibiting infection of a
XX CD4+ cell by HIV. The method comprises contacting the HIV with a CD4-IgG2
XX chimeric heterotetramer to form a complex with the HIV (CD4 = cluster of
XX differentiation 4; IgG2 = immunoglobulin gamma 2). CD4 is a cell surface
XX glycoprotein that is expressed primarily on the surface of T cells. In
XX man, CD4 is the target of interaction with HIV. The heterotetramer has
XX two heavy and two light chains which are encoded by expression vectors
XX CD4-IgG2HC-PRCMV (VI) and CD4-kLC-PRCMV (V2), respectively. The method
XX is used to inhibit infection of a CD4+ cell by a HIV and to prevent CD4+
XX cells of a subject from becoming infected with HIV. The method is also
XX useful for treating a subject having CD4+ cells infected with HIV. The
XX present sequence is human fusion protein: CD4-Ig2 chimeric heavy chain of
XX the CD4-IgG2 chimeric heterotetramer. This sequence was used in the
XX method of the present invention
XX
XX Sequence 530 AA;
XX
XX Query Match 62.0%; Score 2116; DB 4; Length 530;
XX Best Local Similarity 68.9%; Pred. No. 2.8e-106;
XX Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;
XX

```

```

QY 1 MNRGVPFRHLLVLTQALLPAPATGKNKVLGKGGDTVELTCTASQKSIQPHMKNSQIK 60
DB 1 MNRGVPFRHLLVLTQALLPAPATGKNKVLGKGGDTVELTCTASQKSIQPHMKNSQIK 60
QY 61 ILGNQGSFLTKGPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLQGGSLTTLTSPSSPSVQCRSPRGKNIQGGKTLSTVSOLELDQSG 180
DB 121 LVFGLTANSPTHLQGGSLTTLTSPSSPSVQCRSPRGKNIQGGKTLSTVSOLELDQSG 180
QY 181 TMTCTVLQNGKVEFEKIDIVLAFQKASSIYKKEGEQVEFSFPLAFTVEKLTGSSGELMW 240
DB 181 TMTCTVLQNGKVEFEKIDIVLAFQKASSIYKKEGEQVEFSFPLAFTVEKLTGSSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPRKQNGKKLPLHLTLPQALPOYAG--SGNL 297
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTDPRKQNGKKLPLHLTLPQALPOYAG--SGNL 297
QY 298 TLAEAKTGKLGHOENVLVMPATQL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
DB 298 TLAEAKTGKLGHOENVLVMPATQL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
QY 357 PVMVLNPEAGMWQCLSDSGVLLSENIKVLPTWSTPV-----PCPAPEPKSCDKHTCP 411
DB 357 PVMVLNPEAGMWQCLSDSGVLLSENIKVLPTWSTPV-----PCPAPEPKSCDKHTCP 411
QY 412 ELGQPSVFLFPPPKKDTLMISRTPEVTCVAVDVSHEDPEVKFMVYDGVENVNAKTKPR 471
DB 412 ELGQPSVFLFPPPKKDTLMISRTPEVTCVAVDVSHEDPEVKFMVYDGVENVNAKTKPR 471
QY 472 EEQYNSTFRVSVLTVLHODMLNGKEYCKVSNKALPAPIKTIISKAKGQPREQVYTLR 531
DB 472 EEQYNSTFRVSVLTVLHODMLNGKEYCKVSNKALPAPIKTIISKAKGQPREQVYTLR 531
QY 532 PSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTTPVLDSDGSFPLYSKLT 591
DB 532 PSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTTPVLDSDGSFPLYSKLT 591
QY 592 DKSRMQQGNVSCSVMEALHNHYTQKSLSLSPG 625
DB 592 DKSRMQQGNVSCSVMEALHNHYTQKSLSLSPG 625
QY 496 DKSRMQQGNVSCSVMEALHNHYTQKSLSLSPG 529
DB 496 DKSRMQQGNVSCSVMEALHNHYTQKSLSLSPG 529

RESULT 19
ABG71123 standard; protein; 530 AA.
XX ABG71123;
AC ABG71123;
XX
DT 17-JAN-2003 (first entry)
XX
DE CD4-immunoglobulin G2 (IgG2) chimeric heterotetramer.
XX
KM CD4; immunoglobulin G2; Ig gamma2; human immunodeficiency virus-1; HIV-1;
KM mutant; mltcin.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..25
FT Protein /label=Signal_peptide
FT 26..530
FT /note="Mature CD4-IgG2 chimeric heterotetramer"
XX
PN US6451313-B1.
XX
PD 17-SEP-2002.
XX
PF 07-JUN-1995; 95US-00484681.

```

```

XX 08-FEB-1991; 91US-00653684.
PR 10-FEB-1992; 92MO-US001143.
PR 08-DEC-1992; 92US-00960440.
XX
XX (PROG-) PROGENICS PHARM INC.
PI Maddon PJ, Beaudry GA;
DR N-PSDB; ABS55721.
DR WPI; 2003-038273/03.
XX
XX Novel CD4-immunoglobulin G2 chimeric heterotetramer neutralizes human
PT immunodeficiency virus-1 with two heavy and light chains encoded by
PT expression vectors designated CD4-IgG2HC-prcCMV and CD4-KLC-prcCMV,
PT respectively.
XX
XX Claim 1; Fig 4A-H; 54p; English.
XX
XX The invention describes a purified CD4-immunoglobulin (Ig)G2 chimeric
CC heterotetramer (I) that neutralises human immunodeficiency virus-1 (HIV-
CC 1) having two heavy chains encoded by an expression vector designated CD4
CC -IgG2HC-prcCMV, and two light chains encoded by expression vector
CC designated CD4-KLC-prcCMV. (I) and a composition (II) comprising (I) or
CC (I) linked to a toxin, are useful for inhibiting HIV infection of a CD4
CC cell, and preventing a subject being infected with HIV by blocking the
CC spread of HIV infection. This is the amino acid sequence of the CD4-
CC immunoglobulin G2 chimeric heavy chain chimeric heterotetramer useful in
CC inhibiting HIV infection
XX
XX Sequence 530 AA:
SO
Query Match 62.0%; Score 2116; DB 6; Length 530;
Best Local Similarity 68.9%; Pred. No. 2,86-106;
Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

QY 1 MNRGVPFRHLLVLTQALLPAPATGKNKVLGKGGDTVELTCTASQKSIQPHMKNSQIK 60
DB 1 MNRGVPFRHLLVLTQALLPAPATGKNKVLGKGGDTVELTCTASQKSIQPHMKNSQIK 60
QY 61 ILGNQGSFLTKGPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLQGGSLTTLTSPSSPSVQCRSPRGKNIQGGKTLSTVSOLELDQSG 180
DB 121 LVFGLTANSPTHLQGGSLTTLTSPSSPSVQCRSPRGKNIQGGKTLSTVSOLELDQSG 180
QY 181 TMTCTVLQNGKVEFEKIDIVLAFQKASSIYKKEGEQVEFSFPLAFTVEKLTGSSGELMW 240
DB 181 TMTCTVLQNGKVEFEKIDIVLAFQKASSIYKKEGEQVEFSFPLAFTVEKLTGSSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPRKQNGKKLPLHLTLPQALPOYAG--SGNL 297
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTDPRKQNGKKLPLHLTLPQALPOYAG--SGNL 297
QY 298 TLAEAKTGKLGHOENVLVMPATQL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
DB 298 TLAEAKTGKLGHOENVLVMPATQL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
QY 357 PVMVLNPEAGMWQCLSDSGVLLSENIKVLPTWSTPV-----PCPAPEPKSCDKHTCP 411
DB 357 PVMVLNPEAGMWQCLSDSGVLLSENIKVLPTWSTPV-----PCPAPEPKSCDKHTCP 411
QY 412 ELGQPSVFLFPPPKKDTLMISRTPEVTCVAVDVSHEDPEVKFMVYDGVENVNAKTKPR 471
DB 412 ELGQPSVFLFPPPKKDTLMISRTPEVTCVAVDVSHEDPEVKFMVYDGVENVNAKTKPR 471
QY 472 EEQYNSTFRVSVLTVLHODMLNGKEYCKVSNKALPAPIKTIISKAKGQPREQVYTLR 531
DB 472 EEQYNSTFRVSVLTVLHODMLNGKEYCKVSNKALPAPIKTIISKAKGQPREQVYTLR 531
QY 532 PSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTTPVLDSDGSFPLYSKLT 591
DB 532 PSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTTPVLDSDGSFPLYSKLT 591

```

```

Db      436 PSSEETKQVSLTCLVKGFPSPDIAVEMESNQPENNYKTPMLDSGSPFLYSKLTIV 495
QY      592 DKSRMOQGNVSCSVNHEALHNHYTOKSLSPG 625
Db      496 DKSRMOQGNVSCSVNHEALHNHYTOKSLSPG 529

RESULT 20
AAE37574
ID      AAE37574 standard; protein; 450 AA.
XX
XX      AAE37574;
AC
XX      27-AUG-2003 (first entry)
DT
XX      Human D1D2-Ig alphaCp fusion protein.
DE
XX      Human; CD4; cluster of differentiation factor 4; immunoglobulin; Ig;
KM      human immunodeficiency virus; gene therapy; vaccine; HIV-1 infection; D1;
KW      D2; alpha tailpiece; alphaCp; fusion protein.
XX
XX      Homo sapiens.
OS
XX      MO2003040311-A2.
PN
XX      15-MAY-2003.
PD
XX      24-OCT-2002; 2002MO-US034393.
PF
XX      25-OCT-2001; 2001US-0346231P.
PR
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX      Archos J, Cicala C, Fauci AS;
XX      WPI; 2003-441545/41.
XX      DR      N-PSDB; AAD29113.
XX      DR      N-PSDB; AAD29113.
XX      PT      New CD4 polypeptide ligated at its C-terminus with a portion of an
XX      PT      immunoglobulin, useful for preparing a composition for treating or
XX      PT      preventing HIV-1 infection.
XX      PS      Example 1; Page 47; 100pp; English.
XX      PS      The invention relates to a CD4 (cluster of differentiation factor 4)
XX      CC      polypeptide ligated at its C-terminus with a portion of an immunoglobulin
XX      CC      (Ig) comprising a hinge region and a constant domain of a mammalian Ig
XX      CC      heavy chain. The polypeptide comprises a tailpiece from the C-terminus of
XX      CC      the heavy chain of an IgA or IgM antibody. Polypeptides of the invention
XX      CC      are useful for preparing a composition for treating or preventing human
XX      CC      immunodeficiency virus (HIV)-1 infection. The invention is useful in gene
XX      CC      therapy and also in the preparation of vaccines. The present sequence is
XX      CC      a fusion protein which comprises a human IgA alpha tailpiece (alphaCp), a
XX      CC      human IgG constant region comprising a hinge, a CH2 and CH3 region and a
XX      CC      human CD4 D1D2 domain
XX      SQ      Sequence 450 AA;
SQ
Query Match      61.7%; Score 2107.5; DB 6; Length 450;
Best Local Similarity 65.7%; Pred. No. 6.9e-106;
Matches 426; Conservative 2; Mismatches 5; Indels 215; Gaps 3;
QY      1 MNRGVPFRHLLLVQLALPAPATQGNKRVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Db      1 MNRGVPFRHLLLVQLALPAPATQGNKRVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
QY      61 ILNQGSLFKGSKLNDPRDSRSLSMDQGNFPLITNKLIEDSPYICVEDEQKEVOL 120
Db      61 ILNQGSLFKGSKLNDPRDSRSLSMDQGNFPLITNKLIEDSPYICVEDEQKEVOL 120
QY      121 LVFGLTANSPTHLLOQGSFLTLESPPGSSPSVQCRPRGKNIQGGKTLVSQLELDQSG 180

```

```

Db      121 LVFGLTANSPTHLLOQGSFLTLESPPGSSPSVQCRPRGKNIQGGKTLVSQLELDQSG 180
QY      181 TWTCTVLQNKQKVEKIDIVLAFQKASSIVYKKEGEQVEFPPLAFYBKLTGSGELMW 240
Db      181 TWTCTVLQNKQKVEKIDIVLAFQKASSIVYKKEGEQVEFPPLAFYBKLTGSGELMW 240
QY      241 QABRASSKSWITTFDLKNEVSVKRVTDQPKLQWQKPLPLHLTPQALPQVAGSGLTLA 300
Db      204 -----
QY      301 LEAKTKLHQEVNLVVMRATOLQKNLTCBWGPTSPKMLSLKENKAKVSRKRPVW 360
Db      204 -----
QY      361 LNPBAGMOCCLSDSGQVILBSNIVLPTWSTPVCBPAPBPKSCDKTTC-----PELIG 415
Db      204 -----SADKTHCPCPCAPPELLG 221
QY      416 GPSVFLFPKPKXDTLMISRTPEVTCVVDVSHEDPEVKFNNYVDGVVHNAKTKPREEQY 475
Db      222 GPSVFLFPKPKXDTLMISRTPEVTCVVDVSHEDPEVKFNNYVDGVVHNAKTKPREEQY 281
QY      476 NSTYRVSVTLVTHQDMLNGKEYKCKVSNKALPAPIEKTISKAKGPRBPQVYTLPPSPD 535
Db      282 NSTYRVSVTLVTHQDMLNGKEYKCKVSNKALPAPIEKTISKAKGPRBPQVYTLPPSPD 341
QY      536 ELTKQVSLTCLVKGFPSPDIAVEMESNQPENNYKTPPVLDSDGSFFLYSKLTVDKSR 595
Db      342 ELTKQVSLTCLVKGFPSPDIAVEMESNQPENNYKTPPVLDSDGSFFLYSKLTVDKSR 401
QY      596 MOQGNVSCSVNHEALHNHYTOKSLSPG-----LQDDETC 632
Db      402 MOQGNVSCSVNHEALHNHYTOKSLSPG-----LQDDETC 449

RESULT 21
AAR46679
ID      AAR46679 standard; protein; 530 AA.
XX
XX      AC      AAR46679;
XX      AC      AAR46679;
XX      DT      25-MAR-2003 (revised)
XX      DT      08-AUG-1994 (first entry)
XX      DE      CD4-IgG2 chimeric heavy chain.
XX      KW      CD4; gamma; heavy chain; chimeric; chimaeric; immunconjugate; HIV;
XX      KW      human immunodeficiency virus; radionuclide; toxin; therapy; treatment;
XX      KW      imaging; detection; targeting; immunoglobulin; IgG.
XX      OS      Homo sapiens.
XX      FH      Key
XX      FH      Location/Qualifiers
XX      FT      Region
XX      FT      /label= CD4 Region.
XX      FT      205..302
XX      FT      Region
XX      FT      /label= CH1 Region.
XX      FT      303..314
XX      FT      Region
XX      FT      /label= Hinge Region.
XX      FT      315..423
XX      FT      /label= CH2 Region.
XX      FT      424..530
XX      FT      Region
XX      FT      /label= CH3 Region.
XX      PN      MO9403191-A1.
XX      PD      17-FEB-1994.
XX      PD      17-FEB-1994.
XX      PF      06-AUG-1993; 93MO-US007422.
XX      PR      07-AUG-1992; 92US-00927931.
XX      PA      (PROG-) PROGENICS PHARM INC.

```


XX Allaway GP, Maddon PJ;
 PI WPI: 1994-065392/08.
 DR N-PSDB; MA055751.
 DR Non-peptidyl toxin or radionuclide and CD4-gamma 2 or CD4-IgG2
 PT immunoconjugates - used to kill HIV-infected cells and to image and
 PT stage HIV infection.
 XX
 PS Disclosure; Fig 4; 142pp; English.
 XX
 CC A tetramer comprising 2 IgG2 heavy chains or two CD4-IgG2 chimeric heavy
 CC chains and two kappa light chains or CD4-kappa light chains (AAR4680)
 CC linked to a non-peptidyl toxin or a gamma radiation-emitting radionuclide
 CC of low to moderate cytotoxicity. The resulting immunoconjugate comprising
 CC the toxin can be used to kill HIV infected cells and to treat HIV
 CC infected subjects to reduce the population of HIV infected cells. It can
 CC also be used to reduce the likelihood of infection. The immunoconjugate
 CC comprising the radionuclide can be used to image HIV infected tissue, to
 CC calculate the stage of HIV infection or the efficacy of an anti-HIV
 CC treatment using the imaging technique and for determining the prognosis
 CC of an HIV infected subject. (updated on 25-Mar-2003 to correct PN field.)
 CC
 XX
 SQ Sequence 530 AA;
 Query Match 61.7%; Score 2106; DB 2; Length 530;
 Best Local Similarity 68.5%; Pred. No. 9.8e-106;
 Matches 434; Conservative 28; Mismatches 58; Indels 114; Gaps 10;
 QY 1 MNRGVPFRHLLVQLALPAATOGKRVLGKGDVVELTCTASOKKSIOFHKNSNOIK 60
 DB 1 MNRGVPFRHLLVQLALPAATOGKRVLGKGDVVELTCTASOKKSIOFHKNSNOIK 60
 QY 61 ILGNQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
 DB 61 ILGNRGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
 QY 121 LVFGLTNSDPHLLQGSGLTTLTSPGSSPVOCRRPRGNITGGKTLSTVSOELDDSG 180
 DB 121 LVFGLTNSDPHLLQGSGLTTLTSPGSSPVOCRRPRGNITGGKTLSTVSOELDDSG 180
 QY 181 TWFTVLONOKKVEFKIDIVLAFQKASSIYKKEGQVEFSPFLAFTVEKLTGSGELMW 240
 DB 181 TWFTVLONOKKVEFKIDIVLAF--AST-----KQSV--PFLA----- 216
 QY 241 QAERASSKSWITPDLKNKEVSVKRVTDQPKLQMGKTLPHLTLPLQALPOYAG--SGNL 297
 DB 217 -----PCSRSTSESTALAGCLVKQYFPPVTVSNMSGALTSQVH 255
 QY 298 TLAEAKTKLHGVNLVVMRATOL-QKULTCEVWGFTSPKMLSLKENKAVSKREK 356
 DB 256 TFPVLTSSGLYSLSVTVTPSSNFGTQTYCNV-----DHK 292
 QY 357 PWWVLNDEAGWMOGLSDSGVLLSESNIKVLTPTWSTPV-----PCPAPEPKSCDKHTCP 411
 DB 293 P-----SNTKVDKYTEKCCVCEPCPPAP----- 317
 QY 412 ELLEGPSVFLPPKPKDTLMISRTPEYTCVVDVSHEDPEVKNMYVDGVEVNAKTKPR 471
 DB 318 --VAGPSPVFLPPKPKDTLMISRTPEYTCVVDVSHEDPEVKNMYVDGVEVNAKTKPR 375
 QY 472 EEOYNSTYRVVSVLTIVHOMLNGKEYKCKVSKNKAAPLEKITSKAKGQPREPQVYTLR 531
 DB 376 EEOYNSTYRVVSVLTIVHOMLNGKQYCKVSKNKAAPLEKITSKAKGQPREPQVYTLR 435
 QY 532 PSRELTKNQVSLTCLVKGFYPSDIAVEMESNGOPEKNYKTPPVLDSDGSFFLYSKLTIV 591
 DB 436 PSREMTKMQSLTCLVKGFYPSDIAVEMESNGOPEKNYKTPPVLDSDGSFFLYSKLTIV 495
 QY 592 DKSRMOQGNVSCSVMEALHNHYTOKSLSLSPG 625
 DB 496 DKSRMOQGNVSCSVMEALHNHYTOKSLSLSPG 529

RESULT 22
 AAE37576
 ID AAE37576 standard; protein; 449 AA.
 XX
 AC AAE37576;
 XX
 DT 27-AUG-2003 (first entry)
 XX
 DE Human FDI2-Ig alphatp fusion protein variant.
 XX
 KW Human; CD4; cluster of differentiation factor 4; immunoglobulin; Ig;
 KW human immunodeficiency virus; gene therapy; vaccine; HIV-1 infection; D1;
 KW D2; alpha tailpiece; alphatp; fusion protein; mutein; variant; mutant.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 218 /note= "Wild type Glu substituted with Pro"
 FT Misc-difference 219 /note= "Wild type Leu substituted with Val"
 FT Misc-difference 221 /note= "Wild type Gly substituted with Ala"
 FT
 XX W02003040311-A2.
 XX
 PD 15-MAY-2003.
 XX
 PF 24-OCT-2002; 2002WO-US034393.
 XX
 PR 25-OCT-2001; 2001US-0346231P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Athoe J, Cicala C, Fauci AS;
 XX
 DR WPI: 2003-441545/41.
 DR N-PSDB; ACC62877.
 XX
 PT New CD4 polypeptide ligated at its C-terminus with a portion of an
 PT immunoglobulin; useful for preparing a composition for treating or
 PT preventing HIV-1 infection.
 XX
 PS Example 11; Page 67; 100pp; English.
 XX
 CC The invention relates to a CD4 (cluster of differentiation factor 4)
 CC polypeptide ligated at its C-terminus with a portion of an immunoglobulin
 CC (Ig) comprising a hinge region and a constant domain of a mammalian Ig
 CC heavy chain. The polypeptide comprises a tailpiece from the C-terminus of
 CC the heavy chain of an IgA or IgM antibody. Polypeptides of the invention
 CC are useful for preparing a composition for treating or preventing human
 CC immunodeficiency virus (HIV)-1 infection. The invention is useful in gene
 CC therapy and also in the preparation of vaccines. The present sequence is
 CC a fusion protein variant (G218P/L219V/220delA/G221A) which comprises a
 CC human IgA alpha tailpiece (alphatp), a human IgG2 constant region
 CC comprising a hinge, a CH2 and CH3 region and a human CD4 DID2 domain.
 CC This variant protein is also referred to as mutant F
 XX
 SQ Sequence 449 AA;
 Query Match 61.2%; Score 2088; DB 6; Length 449;
 Best Local Similarity 65.2%; Pred. No. 7.7e-105;
 Matches 422; Conservative 3; Mismatches 8; Indels 214; Gaps 3;
 QY 1 MNRGVPFRHLLVQLALPAATOGKRVLGKGDVVELTCTASOKKSIOFHKNSNOIK 60
 DB 1 MNRGVPFRHLLVQLALPAATOGKRVLGKGDVVELTCTASOKKSIOFHKNSNOIK 60
 QY 61 ILGNQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
 DB 61 ILGNQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120

```

QY 121 LVFGLTANSDTHLLOGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLLEIYQDSG 180
DB 121 LVFGLTANSDTHLLOGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLLEIYQDSG 180
QY 181 TWTCVTLQONOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPFLAFVTEKLTGSGELMW 240
DB 181 TWTCVTLQONOKKVEFKIDIVLA----- 203
QY 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPLPALPOYAGSGNLTLLA 300
DB 204 ----- 203
QY 301 LEAKTGKLEHGVNLVYMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWV 360
DB 204 ----- 203
QY 361 LNPEAGMOCCLSDSGQVLLLESNIKVLPTWSTVPCBPAPPEKSCDKHTTCPEL---LGG 416
DB 204 -----SADKTHTCPCBPAPPVAG 221
QY 417 PSVFLPPPKKDTLMISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTKPREEQYN 476
DB 222 PSVFLPPPKKDTLMISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTKPREEQYN 281
QY 477 STYRVSVVLTVLHQMNLNGEKYCKVSNKALPAPIEKITSKAGQPREPQVYTLPPSRDE 536
DB 282 STYRVSVVLTVLHQMNLNGEKYCKVSNKALPAPIEKITSKAGQPREPQVYTLPPSRDE 341
QY 537 LTNGVSLTCLVGFYPSDIAVEMESNGQPENNYKTPPLVLDGSGFFLYSKLTVYKSNW 596
DB 342 LTNGVSLTCLVGFYPSDIAVEMESNGQPENNYKTPPLVLDGSGFFLYSKLTVYKSNW 401
QY 597 QCGNVFSCSVMEALHNHYTQKSLSPG-----LQDETC 632
DB 402 QCGNVFSCSVMEALHNHYTQKSLSPGKPTGVNVSVMVAEVDGTC 448

```

RESULT 23

AAB19510 ID AAB19510 standard; protein; 481 AA.

AAB19510;

09-JAN-2001 (first entry)

CD4-IgM fusion protein CH4Pmu.

CD4; IgM; human; CD4Pmu; fusion protein; immunoglobulin; HIV; SIV; gp120;

therapy; diagnosis.

Homo sapiens.

Key Location/Qualifiers

FH Protein 1..395

FT Protein /note="CD4 extracellular region"

FT Protein /note="IgM heavy chain partial sequence"

US6117656-A.

12-SEP-2000.

07-JUN-1995; 95US-00479353.

22-JAN-1988; 88US-00147351.

09-JUN-1989; 89US-00299596.

12-APR-1993; 93US-00057952.

04-FEB-1994; 94US-00191708.

(GEHO) GEN HOSPITAL CORP.

```

PI Seed B;
XX WPI; 2000-586558/55.
DR N-PSDB; AAA50662.
PT CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or
PT SIV.
XX Example 1; Col 49-60; 39pp; English.
PS
XX The present sequence is that of fusion protein CD4Pmu comprising the
XX extracellular portion of CD4, which binds to HIV gp120, linked at its C-
XX terminus to the human IgM heavy chain. To obtain the fusion protein, DNA
XX encoding CD4 was linked to IgM DNA at the Pst site upstream of the CH2
XX region (see AAA50663). Fusion protein CD4Pmu and a nucleic acid encoding
XX it are claimed. Also claimed are a vector comprising the nucleic acid,
XX and a method of producing the fusion protein in secreted form using a
XX transformed host cell. The fusion protein may further comprise a
XX therapeutic agent, radiolabel or NMR imaging agent. The fusion protein
XX can be administered to an animal (including humans) for treatment of HIV
XX or SIV infection, and can also be used in assays for HIV or SIV, imaging
XX and tissue stains. IgM fusion proteins such as CD4Pmu provide complement-
XX mediated immunity
SQ Sequence 481 AA;

```

Query Match 60.7%; Score 2073; DB 3; Length 481;

Best Local Similarity 84.0%; Pred. No. 5,4e-104; Indels 24; Gaps 6;

Matches 416; Conservative 19; Mismatches 36; Indels 24; Gaps 6;

```

QY 1 NMRGVFRLHLVLOLALLPAATQGNKVLGKGGPTVELTCTASQKSIQFMKNSNQIK 60
DB 1 NMRGVFRLHLVLOLALLPAATQGNKVLGKGGPTVELTCTASQKSIQFMKNSNQIK 60
QY 61 ILGNQSFLLTKGPSKLNDRADSRSLMDQGNPFLIINKLIDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQSFLLTKGPSKLNDRADSRSLMDQGNPFLIINKLIDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSDTHLLOGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLLEIYQDSG 180
DB 121 LVFGLTANSDTHLLOGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLLEIYQDSG 180
QY 181 TWTCVTLQONOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPFLAFVTEKLTGSGELMW 240
DB 181 TWTCVTLQONOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPFLAFVTEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPLPALPOYAGSGNLTLLA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPLPALPOYAGSGNLTLLA 300
QY 301 LEAKTGKLEHGVNLVYMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWV 360
DB 301 LEAKTGKLEHGVNLVYMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWV 360
QY 361 LNPEAGMOCCLSDSGQVLLLESNIKVLPTWSTVPCBPAPPEKSCDKHTTCPELGG--GPS 418
DB 361 LNPEAGMOCCLSDSGQVLLLESNIKVLPTWSTVPCBPAPPEKSCDKHTTCPELGG--GPS 418
QY 419 VFLPPPKKDTLMISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTKPRE----E 472
DB 408 VGVFVPRDGFNGNPKSKLTCQATGFSR--QIQVSWLRGKQVGSVTTDQVQAEAK 464
QY 473 EQNSTYRVSVVLTIV 487
DB 465 ESGPTTYKVTSTLTI 479

```

RESULT 24

AA559171 ID AA559171 standard; protein; 481 AA.

AA559171;

XX

```
DT 14-MAR-2000 (first entry)
XX
XX CD4-Ig fusion protein CD4Pmu.
DE
KM HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;
KM secreted protein; SIV infection; medicament.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX CA1340741-C.
XX
XX 14-SEP-1999.
XX
XX 20-JAN-1989; 89CA-00588749.
XX
XX 20-JAN-1989; 89CA-00588749.
XX
XX (GEO ) GEN HOSPITAL CORP.
XX
XX Seed B;
XX
XX WPI; 2000-063015/06.
XX
XX N-PSDB; AAZ48204.
XX
XX
XX New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in
XX the treatment of HIV or simian immunodeficiency virus infections.
XX
XX Example 1; Page 54-60; 89pp; English.
XX
XX The invention provides a fusion gene encoding a fusion protein that
XX comprises an extracellular CD4 DNA sequence or its fragment which binds
XX to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA
XX sequence of an Ig heavy or light chain, where the DNA sequence encoding
XX the variable region has been replaced with the DNA sequence which encodes
XX extracellular CD4 or its gp120 binding fragment. The fusion protein is
XX capable of being secreted. The fusion proteins are useful for treating
XX HIV or SIV infections in animals, preferably humans. They are also useful
XX for producing medicaments which can be used for treating HIV or SIV
XX infections in humans. The present sequence represents the fusion protein
XX CD4Pmu where the CD4 is linked to human IgG1 at the Fc site upstream of
XX the CH2 region
XX
XX Sequence 481 AA;
SQ
Query Match 60.7%; Score 2073; DB 3; Length 481;
Best Local Similarity 84.0%; Pred. No. 5.4e-104;
Matches 416; Conservative 19; Mismatches 36; Indels 24; Gaps 6;
QY 1 MNRGVPRHLLVQLALPAATQGNKVYLGKGDVVELTCTASOKSIOFHMKNQIK 60
DB 1 MNRGVPRHLLVQLALPAATQGNKVYLGKGDVVELTCTASOKSIOFHMKNQIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLIQGSFLTLESPPSSSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
DB 121 LVFGLTANSPTHLIQGSFLTLESPPSSSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
QY 121 LVFGLTANSPTHLIQGSFLTLESPPSSSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
DB 121 LVFGLTANSPTHLIQGSFLTLESPPSSSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
QY 181 TWTCTVONONKKEVEFKIDIVLAFOKASIVYKKEGEVVEFPALFVEKLTGSGELMW 240
DB 181 TWTCTVONONKKEVEFKIDIVLAFOKASIVYKKEGEVVEFPALFVEKLTGSGELMW 240
QY 241 QAEBASSSKSWITFDLKNKEVSVKRVTPDPLQMGKLLPLHLTPQALPQYAGSGNTLIA 300
DB 241 QAEBASSSKSWITFDLKNKEVSVKRVTPDPLQMGKLLPLHLTPQALPQYAGSGNTLIA 300
QY 301 LEATTGKLHQEVNLVVRATOLQKLTCEVWGFTSPKLMSLKLENKEAVYSKREKPVWV 360
DB 301 LEATTGKLHQEVNLVVRATOLQKLTCEVWGFTSPKLMSLKLENKEAVYSKREKPVWV 360
QY 301 LEATTGKLHQEVNLVVRATOLQKLTCEVWGFTSPKLMSLKLENKEAVYSKREKPVWV 360
DB 301 LEATTGKLHQEVNLVVRATOLQKLTCEVWGFTSPKLMSLKLENKEAVYSKREKPVWV 360
```

```
QY 361 LNPEAGMOCILSDSGVLLSESNIKVLPWTSTVPCEPAPBPKSCDKTHTCEBELG--GPS 418
DB 361 LNPEAGMOCILSDSGVLLSESNIKVLPWTSTVP-----HADPEVIAELPRK 407
QY 419 VLFPPPKPQDTLMIS-RTPEVTGVVDVSHEDPEKFNWYDVGVHNAKTKPR----E 472
DB 408 VSVFVP-PRDGFENPRKSKLICQATGFSR--QIVSWLREGKQVSGVTTDQVQAEAK 464
QY 473 EQNSTRYVSVLTV 487
DB 465 ESGPTYKYVSTLTI 479
RESULT 25
AA51081
ID AA51081 standard; protein; 481 AA.
XX
XX AA51081;
XX
XX 23-MAR-2000 (first entry)
XX
XX Human fusion protein CD4Pmu.
XX
XX Fusion protein; human; CD4; IgM; immunoglobulin; gp120;
XX anti-human immunodeficiency virus; CD4Pmu.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX US6004781-A.
XX
XX 21-DEC-1999.
XX
XX 04-FEB-1994; 94US-00191708.
XX
XX 22-JAN-1988; 88US-00147351.
XX
XX 23-JAN-1989; 89US-00299596.
XX
XX 09-JUN-1992; 92US-00896781.
XX
XX 12-APR-1993; 93US-00057952.
XX
XX (GEO ) GEN HOSPITAL CORP.
XX
XX Seed B;
XX
XX WPI; 2000-085792/07.
XX
XX N-PSDB; AAZ44064.
XX
XX Fusion protein useful for the treatment of human immunodeficiency virus.
XX
XX Example 1; Col 49-58; 39pp; English.
XX
XX This invention describes a novel nucleic acid (I) encoding a fusion
XX protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
XX and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
XX chain (III). The products of the invention have anti-human
XX immunodeficiency virus (HIV) activity and are capable of binding to
XX gp120. The fusion protein is useful for treating human immunodeficiency
XX virus (HIV) or simian immunodeficiency virus (SIV). This sequence
XX represents the fusion protein CD4Pmu which is constructed from CD4 linked
XX to human IgM upstream of the CH2 region
XX
XX Sequence 481 AA;
SQ
Query Match 60.5%; Score 2067; DB 3; Length 481;
Best Local Similarity 83.8%; Pred. No. 1.1e-103;
Matches 415; Conservative 18; Mismatches 38; Indels 24; Gaps 6;
QY 1 MNRGVPRHLLVQLALPAATQGNKVYLGKGDVVELTCTASOKSIOFHMKNQIK 60
DB 1 MNRGVPRHLLVQLALPAATQGNKVYLGKGDVVELTCTASOKSIOFHMKNQIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIKLKIEDSDTYICEVEDQKEEVOL 120
```

```

Db      61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDDTHLLQGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKITLSVSGLEIDSG 180
Db      121 LVFGLTANSDDTHLLQGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKITLSVSGLEIDSG 180
Qy      181 TWTCTVLQONQKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPFLAFVTEKLTGSGELMW 240
Db      181 TWTCTVLQONQKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPFLAFVTEKLTGSGELMW 240
Qy      241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLTPQALPOVAGSGNLTLA 300
Db      241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLTPQALPOVAGSGNLTLA 300
Qy      301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEWGPSTSPKMLSLKLENKEKAVSKREKPVWV 360
Db      301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEWGPSTSPKMLSLKLENKEKAVSKREKPVWV 360
Qy      361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPERPKSCDHTHCPELLG--GPS 418
Db      361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPERPKSCDHTHCPELLG--GPS 418
Qy      419 VFLEPPKPKDTLM-ISRTEPVTGVVVDVSHEDPEVKFNMYVDGEVHNAKTKPR-----E 472
Db      408 VSVFVP-PRDGFPGCPKRSKLTICQATGFSR--QIQVSWLRBGKQVGSVTTDQVQAEAK 464
Qy      473 EGYNSTYRVVSULTV 487
Db      465 ESGPTTYKVTSTLTI 479

RESULT 26
AAP3011
ID AAP3011 standard; protein; 481 AA.
XX
AC AAP3011;
XX
DT 25-MAR-2003 (revised)
DT 03-AUG-1992 (first entry)
XX
DE Genetic construct which encodes CD4 linked to human Igm at the Pat site
DE upstream of the CH2 region (fusion protein CD4Pmu).
XX
KM Fusion protein; immunoglobulin-like molecule; HIV; SIV; therapy;
KM diagnosis; CD4; gp120; binding fragment; glycoprotein; variable region.
XX
OS Homo sapiens.
XX
PN EP325262-A.
XX
PD 26-JUL-1989.
XX
PF 20-JAN-1989; 89EP-00100913.
XX
PR 22-JAN-1988; 88US-00147351.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Seed B;
XX
PI WPI; 1989-214472/30.
XX
DR N-PSDB; AAN90359.
XX
PT Immunoglobulin-CD4 fusion proteins - used for treating HIV or SIV
XX infections or detecting HIV or SIV in sample.
XX
PS Example; Table 4, Page 41-47; 68pp; English.
XX
CC The fusion protein genes of the invention pref. comprises CDNA sequences
CC which encode CD4 or a fragment which binds gp120 ligated to an expresion
CC plasmid which encodes an antibody in which the variable region of the
CC gene has been deleted (see M087-02671). The CD4 portion of the fusion
CC protein may comprise the complete CD4 sequence, the 370 AA extracellular

```

```

CC region and the membrane spanning domain, or the extracellular region. The
CC Ig heavy chain is pref. from Igm, IgG1 or IgG3. The following are
CC specifically claimed: fusion proteins CD4H1ambda1, CD4Hmu, CD4Pmu,
CC CD4Iambda1, and CD4Hmu (No. 67608), PCDApIambda (No. 67609) and
CC PCDApIambda1 (No. 67610). (updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 481 AA;
Query Match 60.5%; Score 2065; DB 1; Length 481;
Best Local Similarity 83.8%; Pred. No. 1,4e-103;
Matches 415; Conservative 19; Mismatches 37; Indels 24; Gaps 6;

Qy      1 NMRGVPFRHLVLQALLPATQGNKVYLGKGDVETCTASQKSIQFMKNSNQIK 60
Db      1 NMRGVPFRHLVLQALLPATQGNKVYLGKGDVETCTASQKSIQFMKNSNQIK 60
Qy      61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVOL 120
Db      61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDDTHLLQGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKITLSVSGLEIDSG 180
Db      121 LVFGLTANSDDTHLLQGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKITLSVSGLEIDSG 180
Qy      181 TWTCTVLQONQKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPFLAFVTEKLTGSGELMW 240
Db      181 TWTCTVLQONQKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPFLAFVTEKLTGSGELMW 240
Qy      241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLTPQALPOVAGSGNLTLA 300
Db      241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLTPQALPOVAGSGNLTLA 300
Qy      301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEWGPSTSPKMLSLKLENKEKAVSKREKPVWV 360
Db      301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEWGPSTSPKMLSLKLENKEKAVSKREKPVWV 360
Qy      361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPERPKSCDHTHCPELLG--GPS 418
Db      361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPERPKSCDHTHCPELLG--GPS 418
Qy      419 VFLEPPKPKDTLMIS-RPEEVTGVVVDVSHEDPEVKFNMYVDGEVHNAKTKPR-----E 472
Db      408 VSVFVP-PRDGFPGCPKRSKLTICQATGFSR--QIQVSWLRBGKQVGSVTTDQVQAEAK 464
Qy      473 EGYNSTYRVVSULTV 487
Db      465 ESGPTTYKVTSTLTI 479

RESULT 27
AAR26782
ID AAR26782 standard; protein; 432 AA.
XX
AC AAR26782;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 06-FEB-1993 (first entry)
XX
DE CD4-gamma2 chimeric heavy chain homodimer.
XX
KM homodimer; soluble CD4; T cell receptor; CD4 antigen; high recovery;
KM chimeric; increased serum half life; HIV infection; AIDS; ss.
XX
OS Homo sapiens.
XX
OS Chimeric.
XX
FH Key
FT Domain 1..216
FT Domain /label= CD4
FT Domain 217..325
FT Domain /label= CH2
FT Domain 326..433

```


Sequence 432 AA:

Query Match 60.3%; Score 2059; DB 2; Length 432;
 Best Local Similarity 65.6%; Pred. No. 2.7e-103;
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

QY 1 MNRGVFRRLLVLTQALPAATQGNKVVYGGKGDVTELTCTASQKKSIOFHMKNSNQIK 60
 |||||||
 DB 1 MNRGVFRRLLVLTQALPAATQGNKVVYGGKGDVTELTCTASQKKSIOFHMKNSNQIK 60
 |||||||

QY 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPPLIIKNLKIEDSDTYICEVEDQKEEYVL 120
 |||||||
 DB 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPPLIIKNLKIEDSDTYICEVEDQKEEYVL 120
 |||||||

QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLEIQDSG 180
 |||||||
 DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLEIQDSG 180
 |||||||

QY 181 TWTCVTLQNGKVEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFVTEKLTGSGELMW 240
 |||||||
 DB 181 TWTCVTLQNGKVEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFVTEKLTGSGELMW 240
 |||||||

QY 241 QAEKASSKSWITFDLKNKEVSVKRVYODPKLQMGKKPLHLTLPOLPQYAGSGNLTLLA 300
 |||||||
 DB 241 QAEKASSKSWITFDLKNKEVSVKRVYODPKLQMGKKPLHLTLPOLPQYAGSGNLTLLA 300
 |||||||

QY 301 LEAKTGKLGHOEVLVVMRATQLOKNTLCEVWGPTSPKMLSLKENKAQVSKREKPVWV 360
 |||||||
 DB 301 LEAKTGKLGHOEVLVVMRATQLOKNTLCEVWGPTSPKMLSLKENKAQVSKREKPVWV 360
 |||||||

QY 361 LNPBAGMOCGLSDSGVLLSNIKYLPTMTSPVPCPAPKSCDKHTHCPELLGGPSVF 420
 |||||||
 DB 361 LNPBAGMOCGLSDSGVLLSNIKYLPTMTSPVPCPAPKSCDKHTHCPELLGGPSVF 420
 |||||||

QY 421 LFPFKRDTLMISRTBEVTCVVVDVSHEDPEVKFNNYVDSGEVHNAAKTRREQYNSTYR 480
 |||||||
 DB 421 LFPFKRDTLMISRTBEVTCVVVDVSHEDPEVKFNNYVDSGEVHNAAKTRREQYNSTYR 480
 |||||||

QY 481 VNSVLTVLVHODMNLNGEKYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKN 540
 |||||||
 DB 481 VNSVLTVLVHODMNLNGEKYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKN 540
 |||||||

QY 541 QVSLTCLVNGFVPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRMQQGN 600
 |||||||
 DB 541 QVSLTCLVNGFVPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRMQQGN 600
 |||||||

QY 601 VFSCSVWHEALHNHYTQKSLSLSPG 625
 |||||||
 DB 601 VFSCSVWHEALHNHYTQKSLSLSPG 625
 |||||||

QY 625 VFSCSVWHEALHNHYTQKSLSLSPG 631
 |||||||
 DB 625 VFSCSVWHEALHNHYTQKSLSLSPG 631
 |||||||

RESULT 29
 AAY85079 standard; protein; 432 AA.
 ID AAY85079 standard; protein; 432 AA.
 XX AAY85079;
 AC AAY85079;
 XX 19-JUN-2000 (first entry)
 XX Human CD4-gamma 2 chimeric heavy chain homodimer amino acid sequence.
 DE Human CD4-gamma 2 chimeric heavy chain homodimer amino acid sequence.
 XX CD4-gamma 2 chimeric heavy chain homodimer; immunoglobulin; treatment;
 KM cytotoxic radionuclide; cell surface glycoprotein; prevent; infection;
 KM cellular immune response interaction mediator; HIV interaction; staging;
 KM prognosis; envelope glycoprotein burden; human.
 XX Homo sapiens.
 OS Homo sapiens.
 PN US6034223-A.
 XX 07-MAR-2000.
 PD 07-MAR-2000.
 XX 07-JUN-1995; 95US-00477460.

XX 07-AUG-1992; 92US-00927931.
 PR 06-AUG-1993; 93MO-US007422.
 PR 03-FEB-1995; 95US-00379516.
 XX (PROG-) PROGENICS PHARM INC.
 PA Allaway GP, Madden PJ;
 PI WPI; 2000-269502/23.
 DR N-PSDB; AA298855.
 XX New immunoglobulin, used to treat, prevent or image human immune
 PT deficiency virus infection, comprises radionuclide attached to
 PS heterodimer of CD4-immunoglobulin chimeras.
 XX Disclosure; Fig 3; 58p; English.
 XX This sequence represents the human CD4-gamma 2 chimeric heavy chain
 CC homodimer amino acid sequence. The invention relates to an
 CC immunoglobulin comprising a cytotoxic radionuclide and a heterodimer
 CC of two heavy chains and two light chains. The cytotoxic radionuclide is
 CC linked to either the heavy chains or the light chains, or to all four
 CC chains, directly or through a bifunctional chelator. Both heavy chains
 CC are chimeric CD4-Ig (immunoglobulin) G2 chains encoded by vector CD4-
 CC IgG2HC-PRC/MV (ATCC 75193) and both light chains are chimeric CD4-kappa
 CC chains encoded by vector CD4-kLC-PRC/MV (ATCC 75194). CD4 is a non-
 CC polymorphic cell surface glycoprotein that is expressed on the surface of
 CC helper T lymphocytes, cells of the monocyte/macrophage lineage and
 CC dendritic cells. CD4 associates with major histocompatibility complex
 CC (MHC) class II molecules on the surface of antigen presenting cells to
 CC mediate efficient cellular immune response interactions. In humans CD4 is
 CC the target of interaction with the human immunodeficiency virus HIV. The
 CC immunoglobulin is used to kill cells infected with HIV, and for treating
 CC or preventing infection. It is also used for imaging HIV-infected tissues
 CC (for staging or prognosis of infection), and for assessing efficacy of
 CC treatments). The immunoglobulin is also used to determine the HIV
 CC envelope glycoprotein burden, once determined, this information is used
 CC in the staging and prognosis of HIV infected patients. The
 CC immunoglobulin should be active against all strains of HIV (since the
 CC CD4-gp120 interaction is essential for infection). The heterodimers
 CC are assembled intracellularly and secreted efficiently from mammalian
 CC cells, allowing high recovery and purification from the culture medium.
 CC They have longer half-life in serum and greater avidity than heavy chain
 CC dimers
 XX
 XX Sequence 432 AA:
 SQ
 Query Match 60.3%; Score 2059; DB 3; Length 432;
 Best Local Similarity 65.6%; Pred. No. 2.7e-103;
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

QY 1 MNRGVFRRLLVLTQALPAATQGNKVVYGGKGDVTELTCTASQKKSIOFHMKNSNQIK 60
 |||||||
 DB 1 MNRGVFRRLLVLTQALPAATQGNKVVYGGKGDVTELTCTASQKKSIOFHMKNSNQIK 60
 |||||||

QY 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPPLIIKNLKIEDSDTYICEVEDQKEEYVL 120
 |||||||
 DB 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPPLIIKNLKIEDSDTYICEVEDQKEEYVL 120
 |||||||

QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLEIQDSG 180
 |||||||
 DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLEIQDSG 180
 |||||||

QY 181 TWTCVTLQNGKVEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFVTEKLTGSGELMW 240
 |||||||
 DB 181 TWTCVTLQNGKVEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFVTEKLTGSGELMW 240
 |||||||

QY 241 QAEKASSKSWITFDLKNKEVSVKRVYODPKLQMGKKPLHLTLPOLPQYAGSGNLTLLA 300
 |||||||
 DB 241 QAEKASSKSWITFDLKNKEVSVKRVYODPKLQMGKKPLHLTLPOLPQYAGSGNLTLLA 300
 |||||||

QY 301 LEAKTGKLGHOEVLVVMRATQLOKNTLCEVWGPTSPKMLSLKENKAQVSKREKPVWV 360
 |||||||
 DB 301 LEAKTGKLGHOEVLVVMRATQLOKNTLCEVWGPTSPKMLSLKENKAQVSKREKPVWV 360
 |||||||

Db 211 ----- 210
 Qy 361 LNPAGMWQCLSDSGVLLSNIKVLPTWSTPVPCEPAPBPKSCDKHTHTELLGSPVF 420
 Db 211 -----PCCPAP-----VAGPSVF 226
 Qy 421 LFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVDGVEVHNATKPREQYNSTYR 480
 Db 227 LFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVDGVEVHNATKPREQYNSTYR 286
 Qy 481 VVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKN 540
 Db 287 VVSVLTVVHODWLNKGEYKCKVSNKGLPAPIEKTISKAKGQPREPQVYTLPPSRDELTKN 346
 Qy 541 QVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSFPLYSKLTVDKSRMOQGN 600
 Db 347 QVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPMLDSDGSFPLYSKLTVDKSRMOQGN 406
 Qy 601 VFSCSVNHEALHNHYTKSLSPG 625
 Db 407 VFSCSVNHEALHNHYTKSLSPG 431

RESULT 30
 AAB67322
 ID AAB67322 standard; protein; 432 AA.

XX AAB67322;

DT 23-APR-2001 (first entry)

DE CD4-gamma2 chimeric heavy chain homodimer protien.

KM Immunocjugate; chelator; chimeric; HIV; human immunodeficiency virus.

OS Homo sapiens.

PN US6177549-B1.

PD 23-JAN-2001.

PF 10-JUN-1999; 99US-00329916.

PR 07-AUG-1992; 92US-00927931.

PR 06-AUG-1993; 93MO-US007422.

PR 03-FEB-1995; 95US-00379516.

PR 07-JUN-1995; 95US-00477460.

XX (PROG-) PROGENICS PHARM INC.

PI Maddon PJ, Allaway GP;

DR WPI; 2001-158582/16.

XX Immunocjugate for treating human immunodeficiency virus-infected

PT subject, consists of cytotoxic radionuclide linked to heterotetramer

PT comprising two chimeric CD4-IgG2 heavy chains and two chimeric CD4-

PT kappa light chains.

PS Disclosure; Fig 3; 43pp; English.

XX The present invention relates to an immunocjugate, comprising a

XX cytotoxic radionuclide linked, directly or via a bifunctional chelator,

XX to a heterotetramer having two chimeric CD4-IgG2 heavy chains encoded by

XX an expression vector CD4-IgG2HC-PRCMV and two chimeric CD4-kappa light

XX chains encoded by an expression vector CD4-kLC-PRCMV. The invention is

XX useful for killing human immunodeficiency virus (HIV)-infected cells, for

XX the treatment and prevention of infection with HIV

Best Local Similarity 65.6%; Pred. No. 2,7e-103;
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

Qy 1 MNRGVPFRHLLVLTQLALLPAATGAKNYLGKDDTVELTCTASQKSIQFHWKNSQIK 60
 Db 1 MNRGVPFRHLLVLTQLALLPAATGAKNYLGKDDTVELTCTASQKSIQFHWKNSQIK 60
 Qy 61 ILGNQSFLLTKGPKSLMDRDRSRSLMDQGNFLLIKNLKTEDSDTYICEVEDQKEEVOL 120
 Db 61 ILGNQSFLLTKGPKSLMDRDRSRSLMDQGNFLLIKNLKTEDSDTYICEVEDQKEEVOL 120
 Qy 121 LVFGLTANSDTHLLOQSLTTLTSPGSSPSYQCRSPRGKNIQGGKTLVSQLELDSG 180
 Db 121 LVFGLTANSDTHLLOQSLTTLTSPGSSPSYQCRSPRGKNIQGGKTLVSQLELDSG 180
 Qy 181 TWCTTVLQONQKVEFKIDIVYLAFOKASSLYYKKEGQVEBSFPLATVEKLTGSGELMW 240
 Db 181 TWCTTVLQONQKVEFKIDIVYLAFOKASSLYYKKEGQVEBSFPLATVEKLTGSGELMW 240
 Qy 241 QAERASSKSMITFDLKNKEVSVRVTPQDPKLGKRLPLHLTLPQALPOYAGSGNLTIA 300
 Db 241 QAERASSKSMITFDLKNKEVSVRVTPQDPKLGKRLPLHLTLPQALPOYAGSGNLTIA 300
 Qy 301 LEAKTGKLBQBVNLVVRATOLQKILTCFVWGPTSPKILMLIKLENKAKVSKKEKPVWV 360
 Db 211 ----- 210
 Qy 361 LNPAGMWQCLSDSGVLLSNIKVLPTWSTPVPCEPAPBPKSCDKHTHTELLGSPVF 420
 Db 211 -----PCCPAP-----VAGPSVF 226
 Qy 421 LFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVDGVEVHNATKPREQYNSTYR 480
 Db 227 LFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVDGVEVHNATKPREQYNSTYR 286
 Qy 481 VVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKN 540
 Db 287 VVSVLTVVHODWLNKGEYKCKVSNKGLPAPIEKTISKAKGQPREPQVYTLPPSRDELTKN 346
 Qy 541 QVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSFPLYSKLTVDKSRMOQGN 600
 Db 347 QVSLTCLVKGFPYSDIAVEWESNGQPENNYKTTTPMLDSDGSFPLYSKLTVDKSRMOQGN 406
 Qy 601 VFSCSVNHEALHNHYTKSLSPG 625
 Db 407 VFSCSVNHEALHNHYTKSLSPG 431

RESULT 31

AAB80883
 ID AAB80883 standard; protein; 432 AA.

XX AAB80883;

DT 29-MAY-2001 (first entry)

DE Human CD4-gamma2 chimeric heavy chain homodimer.

XX Human; Anti-HIV; CD4-IgG2 chimeric heterotetramer;

XX immunoglobulin gamma 2.

OS Homo sapiens.

PN US6187748-B1.

PD 13-FEB-2001.

PF 07-JUN-1995; 95US-00485372.

PR 08-FEB-1991; 91US-00653684.

PR 10-FEB-1992; 92MO-US001143.

PR 08-DEC-1992; 92US-00960440.

XX Query Match 60.3%; Score 2059; DB 4; Length 432.

PA (PROG-) PROGENICS PHARM INC.
 XX Maddon PJ, Beaudry GA;
 PI
 DR WPI; 2001-264981/27.
 DR N-PSDB; AAF71829.
 XX
 PT Inhibiting human immunodeficiency virus (HIV) infection of a CD4+ cell,
 PT or treating a subject having CD4+ cells infected with HIV involves using
 PT CD4-IgG2 chimeric heterotetramer to form a complex with the HIV.
 XX
 PS Disclosure; Fig 3; 55pp; English.
 XX
 CC The present invention relates to a method for inhibiting infection of a
 CC CD4+ cell by HIV. The method comprises contacting the HIV with a CD4-IgG2
 CC chimeric heterotetramer to form a complex with the HIV (CD4 = cluster of
 CC differentiation 4; IgG2 = immunoglobulin gamma 2). CD4 is a cell surface
 CC glycoprotein that is expressed primarily on the surface of T cells. In
 CC man, CD4 is the target of interaction with HIV. The heterotetramer has
 CC two heavy and two light chains which are encoded by expression vectors
 CC CD4-IgG2HC-PRCMV (VI) and CD4-KLC-PRCMV (V2), respectively. The method
 CC is used to inhibit infection of a CD4+ cell by a HIV and to prevent CD4+
 CC cells of a subject from becoming infected with HIV. The method is also
 CC useful for treating a subject having CD4+ cells infected with HIV. The
 CC present sequence is human fusion protein: CD4-gamma2 chimeric heavy chain
 CC homodimer. This sequence was used in the method of the present invention
 CC
 XX
 XX Sequence 432 AA;
 SQ
 Query Match 60.3%; Score 2059; DB 4; Length 432;
 Best Local Similarity 65.6%; Pred. No. 2,7e-103;
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;
 QY 1 MNRGVFPRHLVLVQLALPAATQGNKVVLGKKGDVVELTCTASOKKSIQFHKNSNQIK 60
 DB 1 MNRGVFPRHLVLVQLALPAATQGNKVVLGKKGDVVELTCTASOKKSIQFHKNSNQIK 60
 QY 61 ILGNQSPFLTKGSPKNDPRADSRSLMDQNFPLIKNKIETDSPTYICVEHQKEVOL 120
 DB 61 ILGNQSPFLTKGSPKNDPRADSRSLMDQNFPLIKNKIETDSPTYICVEHQKEVOL 120
 QY 121 LVEGLTANSRDTLHLLQGOSTLTLESPPGSSPVQCSPPRGKNIQGGKTLSSVQLELDQSG 180
 DB 121 LVEGLTANSRDTLHLLQGOSTLTLESPPGSSPVQCSPPRGKNIQGGKTLSSVQLELDQSG 180
 QY 121 LVEGLTANSRDTLHLLQGOSTLTLESPPGSSPVQCSPPRGKNIQGGKTLSSVQLELDQSG 180
 DB 121 LVEGLTANSRDTLHLLQGOSTLTLESPPGSSPVQCSPPRGKNIQGGKTLSSVQLELDQSG 180
 QY 181 TWTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVEKLTGSGELMW 240
 DB 181 TWTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVEKLTGSGELMW 240
 QY 181 TWTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVEKLTGSGELMW 240
 DB 181 TWTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVEKLTGSGELMW 240
 QY 241 QAERASSKSWITFDLKNKEVSVKRVTPDKLQMGKLPILHLTPQALPOYAGSGNLTLLA 300
 DB 241 QAERASSKSWITFDLKNKEVSVKRVTPDKLQMGKLPILHLTPQALPOYAGSGNLTLLA 300
 QY 211 ----- 210
 DB 211 ----- 210
 QY 301 LEAKTGKLEHQBVLVWRATQLOKNLTCVWGPTSPYMLSLKLENKAKVSREKPVWV 360
 DB 301 LEAKTGKLEHQBVLVWRATQLOKNLTCVWGPTSPYMLSLKLENKAKVSREKPVWV 360
 QY 211 ----- 210
 DB 211 ----- 210
 QY 361 LNPBAGMOCLLSDSGVLLIESNIKVLPTWSTVPCEPAPPKSCDKTHTCPELLGSPVF 420
 DB 361 LNPBAGMOCLLSDSGVLLIESNIKVLPTWSTVPCEPAPPKSCDKTHTCPELLGSPVF 420
 QY 211 -----EC-----PPCPAP-----VAGSVF 226
 DB 211 -----EC-----PPCPAP-----VAGSVF 226
 QY 421 LFPKPKDQTLMTISRTPEVTCVVVDVSHEDPEVKPNWVDVEVHNAAKTRERQVSTVR 480
 DB 421 LFPKPKDQTLMTISRTPEVTCVVVDVSHEDPEVKPNWVDVEVHNAAKTRERQVSTVR 480
 QY 227 LFPKPKDQTLMTISRTPEVTCVVVDVSHEDPEVKPNWVDVEVHNAAKTRERQVSTVR 286
 DB 227 LFPKPKDQTLMTISRTPEVTCVVVDVSHEDPEVKPNWVDVEVHNAAKTRERQVSTVR 286
 QY 481 VNSVLTVLHODMNGEKYKCKVSNKALPAPIEKTIISAKGQPREPQVYTLPPSRDELTKN 540
 DB 481 VNSVLTVLHODMNGEKYKCKVSNKALPAPIEKTIISAKGQPREPQVYTLPPSRDELTKN 540
 QY 287 VNSVLTVLHODMNGEKYKCKVSNKALPAPIEKTIISAKGQPREPQVYTLPPSRDELTKN 346
 DB 287 VNSVLTVLHODMNGEKYKCKVSNKALPAPIEKTIISAKGQPREPQVYTLPPSRDELTKN 346
 QY 541 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTPPVVLDSDGSFPLYSKLTVDXSRWQGN 600
 DB 541 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTPPVVLDSDGSFPLYSKLTVDXSRWQGN 600
 QY 347 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTPPVVLDSDGSFPLYSKLTVDXSRWQGN 406
 DB 347 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTPPVVLDSDGSFPLYSKLTVDXSRWQGN 406

QY 601 VFSCSYMEALAHNYTOKSLSPG 625
 DB 407 VFSCSYMEALAHNYTOKSLSPG 431
 RESULT 32
 ID ABG71122 standard; protein; 432 AA.
 XX ABG71122
 AC ABG71122;
 XX
 DT 17-JAN-2003 (first entry)
 XX
 DE CD4-gamma2 chimeric heavy chain of the CD4-IgG2 chimeric protein.
 XX
 KW CD4; gamma2 heavy chain; human immunodeficiency virus-1; HIV-1; mutant;
 XX
 OS Homo sapiens.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..25
 FT Protein /label=Signal_peptide
 FT 26..432
 FT /note="Mature CD4-gamma2 chimeric heavy chain of the CD4
 FT -IgG2 chimeric heterotetramer"
 XX
 XX US6451313-B1.
 XX
 XX 17-SEP-2002.
 XX
 XX 07-JUN-1995; 95US-00484681.
 XX
 XX 08-FEB-1991; 91US-00653684.
 XX
 XX 10-FEB-1992; 92WO-US001143.
 XX
 XX 08-DEC-1992; 92US-00960440.
 XX
 PA (PROG-) PROGENICS PHARM INC.
 PI Maddon PJ, Beaudry GA;
 DR WPI; 2003-038273/03.
 DR N-PSDB; ABS55720.
 XX
 PT Novel CD4-immunoglobulin G2 chimeric heterotetramer neutralizes human
 PT immunodeficiency virus-1 with two heavy and light chains encoded by
 PT expression vectors designated CD4-IgG2HC-PRCMV and CD4-KLC-PRCMV,
 PT respectively.
 XX
 PS Disclosure; Fig 3A-F; 54pp; English.
 XX
 CC The invention describes a purified CD4-immunoglobulin (Ig)G2 chimeric
 CC heterotetramer (II) that neutralises human immunodeficiency virus-1 (HIV-
 CC 1) having two heavy chains encoded by an expression vector designated CD4
 CC -IGG2HC-PRCMV, and two light chains encoded by expression vector
 CC designated CD4-KLC-PRCMV. (II) and a composition (II) comprising (I) or
 CC (I) linked to a toxin, are useful for inhibiting HIV infection of a CD4
 CC cell, and preventing a subject being infected with HIV by blocking the
 CC spread of HIV infection. This is the amino acid sequence of the CD4-
 CC gamma2 chimeric heavy chain of the CD4-IgG2 chimeric heterotetramer
 CC useful in inhibiting HIV infection
 CC
 XX
 XX Sequence 432 AA;
 SQ
 Query Match 60.3%; Score 2059; DB 6; Length 432;
 Best Local Similarity 65.6%; Pred. No. 2,7e-103;
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;
 QY 1 MNRGVFPRHLVLVQLALPAATQGNKVVLGKKGDVVELTCTASOKKSIQFHKNSNQIK 60
 DB 1 MNRGVFPRHLVLVQLALPAATQGNKVVLGKKGDVVELTCTASOKKSIQFHKNSNQIK 60


```

OY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVL 120
OY 121 LVFGITANSPTHLQGSGLTLTLESPPSSPSVQCRSPRGKNIQGGKTLTSSQLELODSG 180
DB 121 LVFGITANSPTHLQGSGLTLTLESPPSSPSVQCRSPRGKNIQGGKTLTSSQLELODSG 180
OY 181 TWTCTVLQONKQKVEFKIDIVLAFQKASSIYKKKEGEVESPFLAFVETKLTSGGELMW 240
DB 181 TWTCTVLQONKQKVEFKIDIVLAFQKASSIYKKKEGEVESPFLAFVETKLTSGGELMW 240
OY 241 QAEKASSSKSWITFDLNKKEVSVKRVTDPKLQMGKLPPLHLTLPOALPOYAGSGNLTLA 300
DB 241 QAEKASSSKSWITFDLNKKEVSVKRVTDPKLQMGKLPPLHLTLPOALPOYAGSGNLTLA 300
OY 301 LEAKTGKHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
DB 301 LEAKTGKHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
OY 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVPCEPAPPEPKSCDKHTTCEPLLGSPSVF 420
DB 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVPCEPAPPEPKSCDKHTTCEPLLGSPSVF 420
OY 421 LFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKFNMVYDGVENHAKTKPREEQYNSTYR 480
DB 421 LFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKFNMVYDGVENHAKTKPREEQYNSTYR 480
OY 481 VVSUVTLVHOMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKN 540
DB 481 VVSUVTLVHOMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKN 540
OY 541 QVSLTCLVKGFPSPDIAEVESNGQPENNYKTTPEVLDSDGSFLYSKLTVDKSRMGOGN 600
DB 541 QVSLTCLVKGFPSPDIAEVESNGQPENNYKTTPEVLDSDGSFLYSKLTVDKSRMGOGN 600
OY 601 VFSCSVMEALHNHYTQKSLSLSPG 625
DB 601 VFSCSVMEALHNHYTQKSLSLSPG 625
OY 407 VFSCSVMEALHNHYTQKSLSLSPG 431
DB 407 VFSCSVMEALHNHYTQKSLSLSPG 431

RESULT 33
AA51080
ID AA51080 standard; protein; 436 AA.
XX
AC AA51080;
XX
DT 23-MAR-2000 (first entry)
XX
DE Human fusion protein CD4Mg.
XX
KW Fusion protein; human; CD4; Igm; immunoglobulin; gp120;
KW anti-human immunodeficiency virus; CD4Mg.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN US6004781-A.
XX
PD 21-DEC-1999.
XX
PF 04-FEB-1994; 94US-00191708.
XX
PR 22-JAN-1988; 88US-00147351.
PR 23-JUN-1989; 89US-00299596.
PR 09-JUN-1992; 92US-00896781.
PR 12-APR-1993; 93US-00057952.
XX
PA (GENO ) GEN HOSPITAL CORP.
XX
PI Seed B;
XX
DR WPI; 2000-085792/07.

```

```

DR N-PSDB; AAZ44063.
PT Fusion protein useful for the treatment of human immunodeficiency virus.
XX
XX Example 1; Col 41-50; 39pp; English.
PS
CC This invention describes a novel nucleic acid (I) encoding a fusion
CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
CC chain (III). The products of the invention have anti-human
CC immunodeficiency virus (HIV) activity and are capable of binding to
CC gp120. The fusion protein is useful for treating human immunodeficiency
CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence
CC represents the fusion protein CD4Mg which is constructed from CD4 linked
CC to human Igm upstream of the CH1 region
XX
SQ Sequence 436 AA;
Query Match 59.9%; Score 2046.5; DB 3; Length 436;
Best Local Similarity 90.4%; Pred. No. 1.3e-102;
Matches 405; Conservative 4; Mismatches 16; Indels 23; Gaps 3;
OY 1 MNRGVPFRHLVLVQLALPAPATQGNRVJLKGKGDVLELTCTASQKKSIOFHWNQNOIK 60
DB 1 MNRGVPFRHLVLVQLALPAPATQGNRVJLKGKGDVLELTCTASQKKSIOFHWNQNOIK 60
OY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVL 120
OY 121 LVFGITANSPTHLQGSGLTLTLESPPSSPSVQCRSPRGKNIQGGKTLTSSQLELODSG 180
DB 121 LVFGITANSPTHLQGSGLTLTLESPPSSPSVQCRSPRGKNIQGGKTLTSSQLELODSG 180
OY 181 TWTCTVLQONKQKVEFKIDIVLAFQKASSIYKKKEGEVESPFLAFVETKLTSGGELMW 240
DB 181 TWTCTVLQONKQKVEFKIDIVLAFQKASSIYKKKEGEVESPFLAFVETKLTSGGELMW 240
OY 241 QAEKASSSKSWITFDLNKKEVSVKRVTDPKLQMGKLPPLHLTLPOALPOYAGSGNLTLA 300
DB 241 QAEKASSSKSWITFDLNKKEVSVKRVTDPKLQMGKLPPLHLTLPOALPOYAGSGNLTLA 300
OY 301 LEAKTGKHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
DB 301 LEAKTGKHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
OY 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVPCEPAPPEPKSCDKHTT-----CPE 412
DB 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVPHADPECPKRPTRKAKLSTTSARTPG 420
OY 413 LLGGPSVFLFPPKPKDITLMSRTPEVTC 440
DB 421 SASAPTLF-----FLVSC 433

RESULT 34
AA59170
ID AA59170 standard; protein; 474 AA.
XX
AC AA59170;
XX
DT 14-MAR-2000 (first entry)
XX
DE CD4-Ig fusion protein CD4Mmu.
XX
KW HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;
KW secreted protein; SIV infection; medicament.
XX
OS Synthetic.
XX
OS Homo sapiens.
XX
PN CA1340741-C.
XX

```

```

14-SEP-1999.
XX DD
XX PF 20-JAN-1989; 89CA-00588749.
XX PR 20-JAN-1989; 89CA-00588749.
XX PA (GEHO ) GEN HOSPITAL CORP.
XX PT Seed B;
XX DR WPI; 2000-063015/06.
XX DR N-PSDB; AAZ48203.
XX XX
XX XX New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in
XX XX the treatment of HIV or simian immunodeficiency virus infections.
XX XX
XX XX Example 1; Page 47-53; 89pp; English.
XX XX
XX XX The invention provides a fusion gene encoding a fusion protein that
XX XX comprises an extracellular CD4 DNA sequence or its fragment which binds
XX XX to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA
XX XX sequence of an Ig heavy or light chain, where the DNA sequence encoding
XX XX the variable region has been replaced with the DNA sequence which encodes
XX XX the extracellular CD4 or its gp120 binding fragment. The fusion protein is
XX XX capable of being secreted. The fusion proteins are useful for treating
XX XX HIV or SIV infections in animals, preferably humans. They are also useful
XX XX for producing medicaments which can be used for treating HIV or SIV
XX XX infections in humans. The present sequence represents the fusion protein
XX XX CD4muu where the CD4 is linked to human IgG1 at the Mac2 site upstream of
XX XX the CH1 region
XX XX
XX XX Sequence 474 AA;
XX XX
XX XX Query Match 59.8%; Score 2041; DB 3; Length 474;
XX XX Best Local Similarity 91.0%; Pred. No. 2.8e-102;
XX XX Matches 406; Conservative 5; Mismatches 31; Indels 4; Gaps 2.
XX XX
XX QY 1 MNRGVPFRLHLVLQALLPAATGQNKVYLGGKGTVELTCTASQKSIQFHWKNSNQIK 60
XX DB 1 MNRGVPFRLHLVLQALLPAATGQNKVYLGGKGTVELTCTASQKSIQFHWKNSNQIK 60
XX QY 1LGNNGSFLTKGPKSLNDRADRSRLMDQGNFPLIINKLIEDSDTYICEVEDQKEEYOL 120
XX DB 61 ILGNNGSFLTKGPKSLNDRADRSRLMDQGNFPLIINKLIEDSDTYICEVEDQKEEYOL 120
XX QY 121 LVFGTJANSDFHLLGGOSLTLTLESPPGSSPVQCRSPRGNKIQGGKTLVSQLELQDSG 180
XX DB 121 LVFGTJANSDFHLLGGOSLTLTLESPPGSSPVQCRSPRGNKIQGGKTLVSQLELQDSG 180
XX QY 181 TWTCTVLONQKVEFKIDIVLAFQKASSIVYKKEGEVSEFPPLAFTYEKLTGSGELMW 240
XX DB 181 TWTCTVLONQKVEFKIDIVLAFQKASSIVYKKEGEVSEFPPLAFTYEKLTGSGELMW 240
XX QY 241 QAERASSSSKSWTTFPLKNKEVSVKRVTDOPKLOMKKKLPILH.TLQALPOVAGSNLTLTA 300
XX DB 241 QAERASSSSKSWTTFPLKNKEVSVKRVTDOPKLOMKKKLPILH.TLQALPOVAGSNLTLTA 300
XX QY 301 LEAKTGLKHQEVNVLVWMRATOLQKNLTCEVWGPTSPKLMLSIKLENKEAKYSKRREKPVWV 360
XX DB 301 LEAKTGLKHQEVNVLVWMRATOLQKNLTCEVWGPTSPKLMLSIKLENKEAKYSKRREKPVWV 360
XX QY 361 INPEAGMWOCLLSDSGQVLLBSNLIKVLPTWSTPVPVPCAPADEPKSCDKTHTCPPLLGGPSVF 420
XX DB 361 INPEAGMWOCLLSDSGQVLLBSNLIKVLPTWSTPVPV---DPILPLTAHPKQGTTLHLSLSD 417
XX QY 421 LFPKPKDTLMSRTPEVTCVVVDVS 446
XX DB 418 TCPKPKTPKAKLS-TFSARTPAADLS 442
XX XX
XX XX RESULT 35
XX XX AAR27278 standard; protein; 532 AA.
XX XX ID AAR27278

```

```

Db      ||| 181 TWCTVLTQNOCKVEFKIDIVLAFQKASSIYKKEGQVEFSFPLAFTVEKLTSSGELMW 240
Qy      ||| 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQWKKLPLHLTLPOALPOYAGSGNLTIA 300
Db      ||| 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQWKKLPLHLTLPOALPOYAGSGNLTIA 300
Qy      ||| 301 LEAKTGKLGHEVNLVWMRATQLOKNLTCEVWGPTSPKMLSLKLENKAKVSKKEKPYWV 360
Db      ||| 301 LEAKTGKLGHEVNLVWMRATQLOKNLTCEVWGPTSPKMLSLKLENKAKVSKKEKPYWV 360
Qy      ||| 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPYPCAPPEKSC 404
Db      ||| 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPYVHA---DPKLC 401

RESULT 36
AAR78678 ID AAR78678 standard; protein; 532 AA.
XX
AC AAR78678;
XX
DT 16-APR-1996 (first entry)
XX
DE T-cell receptor eta.
XX
KW Chimeric receptor; CD4; T-cell receptor eta; HIV; cytolysis;
KM human immunodeficiency virus; adoptive immunotherapy.
XX
OS Homo sapiens.
XX
PN MO9521528-A1.
XX
PD 17-AUG-1995.
XX
PF 12-JAN-1995; 95WO-US000454.
XX
PR 14-FEB-1994; 94US-00195395.
PR 02-AUG-1994; 94US-00284391.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Seed B, Banapour B, Romeo C, Kolanus W;
XX
DR WPI: 1995-292893/38.
XX
DR N-PSDB; AA096124.
XX
PT Target cytolysis of HIV-infected cells - by chimeric CD4 receptor-bearing
PT cells.
XX
PS Example 2; Page 78-79; 118pp; English.
XX
CC Fusion proteins comprising the extracellular domain of CD4 fused to T-
CC cell receptor zeta, gamma or eta (AAR78676-78, respectively) were
CC expressed in CV1 using a vaccine virus vector. These CD4:zeta, CD4:gamma
CC and CD4:eta chimeric receptors mediated cytolysis of targets expressing
CC HIV gp120/41
XX
SQ Sequence 532 AA;

Query Match 59.7%; Score 2036.5; DB 2; Length 532;
Best Local Similarity 96.3%; Pred. No. 5.5e-102;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

```

```

Db      ||| 121 LVFGLTANSPTHLQGGSLTILSPSSVQCRSPRGKNIQGGKTLVSQLELDGSG 180
Qy      ||| 181 TWCTVLTQNOCKVEFKIDIVLAFQKASSIYKKEGQVEFSFPLAFTVEKLTSSGELMW 240
Db      ||| 181 TWCTVLTQNOCKVEFKIDIVLAFQKASSIYKKEGQVEFSFPLAFTVEKLTSSGELMW 240
Qy      ||| 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQWKKLPLHLTLPOALPOYAGSGNLTIA 300
Db      ||| 241 QAERASSSKSWITFDLKNKEVSVKRVTDPLQWKKLPLHLTLPOALPOYAGSGNLTIA 300
Qy      ||| 301 LEAKTGKLGHEVNLVWMRATQLOKNLTCEVWGPTSPKMLSLKLENKAKVSKKEKPYWV 360
Db      ||| 301 LEAKTGKLGHEVNLVWMRATQLOKNLTCEVWGPTSPKMLSLKLENKAKVSKKEKPYWV 360
Qy      ||| 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPYPCAPPEKSC 404
Db      ||| 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPYVHA---DPKLC 401

RESULT 37
AAR89458 ID AAR89458 standard; protein; 532 AA.
XX
AC AAR89458;
XX
DT 26-SEP-1996 (first entry)
XX
DE CD4:eta fusion protein.
XX
KW CD7; transmembrane domain; chimeric receptor; CD5; CD34; CH2; CH3; IgG1;
KM human; CD4; HIV; proteinaceous alpha-helix; T cell; B cell; neutrophil;
KM dendritic cell; therapy; mammal; infection.
XX
OS Synthetic.
XX
PN WO9603883-A1.
XX
PD 15-FEB-1996.
XX
PF 26-JUL-1995; 95WO-US009468.
XX
PR 02-AUG-1994; 94US-00284391.
PR 24-FEB-1995; 95US-00394388.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Seed B, Banapour B, Romeo C, Kolanus W;
XX
DR WPI: 1996-128034/13.
XX
DR N-PSDB; AAT10803.
XX
PT Membrane-bound chimeric receptor comprising extracellular portion
PT including CD4 fragment - cells expressing receptor can be used for
PT treatment of HIV infection.
XX
PS Example 2; Page 80-81; 134pp; English.
XX
CC AAT10801-T10803 represent membrane bound proteinaceous chimeric receptors
CC of the invention. This sequence represents the CD4:eta chimera. The
CC transmembrane region of the chimeric receptor acts to separate the
CC intracellular and extracellular domains of the chimera, and contains a
CC portion of the CD7 (see AAR89440), CD5 or CD34 transmembrane domains.
CC Alternatively, the extracellular portion of the receptor can be separated
CC from the intracellular domain by the hinge, CH2 and CH3 domains of human
CC IgG1 (see AAR89441). The extracellular portion of the chimeric receptor
CC contains a fragment of CD4 (amino acids 1-394 or 1-200 of the CD4
CC sequence, see AAR89450 and AAR89451) which specifically recognises and
CC binds HIV-infected cells, but does not mediate HIV infection. The
CC extracellular domain of the receptor is separated from the cell membrane
CC by 48 or 72 angstroms, or by one or more proteinaceous alpha-helices. The
CC cells expressing the receptor are preferably T cells, B cells,
CC neutrophils, or dendritic cells. The therapeutic cells expressing the

```

```
CC chimeric receptor are administered to a mammal to treat HIV infection
XX
SQ Sequence 532 AA;
Query Match 59.7%; Score 2036.5; DB 2; Length 532;
Best Local Similarity 98.3%; Pred. No. 5.5e-102;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;
QY 1 MNRGVFRRHLVLVQLALPAATQGNKVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVFRRHLVLVQLALPAATQGNKVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60
QY 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPFLIKNLIKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPFLIKNLIKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSRDTLHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQDSG 180
DB 121 LVFGLTANSRDTLHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQDSG 180
QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
QY 241 QARRASSSKSWITFDLKNKEVSVRVYTOPDKLQMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
DB 241 QARRASSSKSWITFDLKNKEVSVRVYTOPDKLQMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
QY 301 LEAKTGLHQBVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSRKRPVWY 360
DB 301 LEAKTGLHQBVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSRKRPVWY 360
QY 361 LNPBAGMOCCLSDSGQVLLSEINIKVLPWTWSTPVHA---DPKLC 401
DB 361 LNPBAGMOCCLSDSGQVLLSEINIKVLPWTWSTPVHA---DPKLC 401
RESULT 38
AAW83141
ID AAW83141 standard; protein; 532 AA.
XX
AAW83141;
AC
XX
DT 03-FEB-1999 (first entry)
XX
DE Chimeric receptor containing human eta polypeptide.
XX
KM Human; zeta; eta; gamma; membrane-bound chimeric receptor; infection;
KM tumour; cancer cell; autoimmune-generated cell; T cell receptor; CD3;
KM CD4; B cell receptor; Fc receptor; pathogen; bacterial; fungal;
KM protozoan; viral.
XX
XX Synthetic.
OS Homo sapiens.
OS
PN US5843728-A.
XX
PD 01-DEC-1998.
XX
PF 05-APR-1995; 95US-00417495.
XX
PR 07-MAR-1991; 91US-00665961.
PR 06-MAR-1992; 92US-00847566.
PR 28-FEB-1994; 94US-00203866.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Romeo C, Kolanus W, Seed B;
XX
DR WPI; 1999-044582/04.
DR N-PSDB; AAV70157.
XX
PT Membrane-bound chimeric receptors - comprising extracellular portion
```

```
PT which recognises and binds a target cell and an intracellular portion of
PT e.g. a T-cell receptor.
XX
PS Claim 11; Col 45-48; 57pp; English.
XX
CC The present invention describes DNA encoding a membrane-bound chimeric
CC receptor comprising: (a) an extracellular portion that specifically
CC recognises and binds a target cell or a target infective agent; and (b)
CC an intracellular portion of a T-cell receptor CD3, zeta or eta
CC polypeptide, a B-cell receptor polypeptide or an Fc receptor polypeptide.
CC The present sequence represents a chimeric receptor containing the human
CC eta polypeptide. Cells expressing chimeric receptors of the present
CC invention can be administered to mammals in order to destroy pathogens
CC (e.g. bacteria, fungi, protozoa or viruses, especially HIV), cancer cells
CC or autoimmune-generated cells
XX
SQ Sequence 532 AA;
Query Match 59.7%; Score 2036.5; DB 2; Length 532;
Best Local Similarity 98.3%; Pred. No. 5.5e-102;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;
QY 1 MNRGVFRRHLVLVQLALPAATQGNKVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVFRRHLVLVQLALPAATQGNKVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60
QY 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPFLIKNLIKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPFLIKNLIKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSRDTLHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQDSG 180
DB 121 LVFGLTANSRDTLHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQDSG 180
QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240
QY 241 QARRASSSKSWITFDLKNKEVSVRVYTOPDKLQMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
DB 241 QARRASSSKSWITFDLKNKEVSVRVYTOPDKLQMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
QY 301 LEAKTGLHQBVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSRKRPVWY 360
DB 301 LEAKTGLHQBVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSRKRPVWY 360
QY 361 LNPBAGMOCCLSDSGQVLLSEINIKVLPWTWSTPVHA---DPKLC 401
DB 361 LNPBAGMOCCLSDSGQVLLSEINIKVLPWTWSTPVHA---DPKLC 401
RESULT 39
AAR27276
ID AAR27276 standard; protein; 575 AA.
XX
AAR27276;
AC
XX
DT 25-MAR-2003 (revised)
DT 28-JUL-1995 (first entry)
XX
DE CD4:zeta peptide chimeric protein.
XX
KM Fusion protein; CD4; extracellular domain; zeta; eta; gamma;
KM membrane spanning domain; intracellular domain; type I;
KM integral membrane homodimer; TCR; T cell antigen receptor;
KM extracellular domain; mouse; human; receptor; chimera;
KM HPB-ALL tumour cell line; natural killer cell.
XX
XX Homo sapiens.
OS
OS
XX
XX Key Location/Qualifiers
XX 1..399
XX FT Protein
XX /note="CD4 extracellular domain"
```